

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 25, 2005, 05:38:23 ; Search time 21022 Seconds  
(without alignments)  
17642.290 Million cell updates/sec

Title: US-10-006-771B-1

Perfect score: 7654

Sequence: 1 aagctgcagctgcaggt.....gtaaaacagcgccagtgcc 7654

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_hcg.\*

3: gb\_in.\*

4: gb\_on.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3666.6	47.9	6501	12 AF132211	AF132211 Cloning v
2	3349	43.8	9320	6 AX663075	AX663075 Sequence
3	3307.8	43.2	6644	6 E23356	E23356 Virus vecto
4	3244.4	42.4	6505	6 AX823827	AX823827 Sequence
5	3223.6	42.1	5903	6 C0879092	C0879092 Sequence
6	3213	42.0	5782	6 AX491314	AX491314 Sequence
7	3183.2	41.6	7372	6 E23357	E23357 Virus vecto
8	3179.2	41.5	6277	12 AB086385	AB086385 Retrovira
9	3140	41.0	6248	12 AB041928	AB041928 Retrovira
10	3130.2	40.9	7257	6 C0879091	C0879091 Sequence
11	3113	40.7	6076	12 AB086386	AB086386 Retrovira
12	3111	40.6	5364	6 AR028670	AR028670 Sequence
13	3104.6	40.6	6506	14 U00220	U00220 Human Immun
14	3096	40.4	6522	6 BD138563	BD138563 Expressio
15	3094.8	40.4	6374	12 SYNMOV1	M64753 Moloney mur
16	3094	40.4	6365	6 I70974	I70974 Sequence 1
17	3087.4	40.3	6824	12 AB086387	AB086387 Retrovira
18	3083.4	40.3	5292	6 AR028669	AR028669 Sequence
19	3083.4	40.3	5294	6 AR028668	AR028668 Sequence

20	3083.4	40.3	5323	6 AR028667	AR028667 Sequence
21	3082.2	40.3	6219	6 BD222120	BD222120 Method an
22	3082.2	40.3	6219	6 AX146810	AX146810 Sequence
23	3062.2	40.0	6145	6 AR140313	AR140313 Sequence
24	3062.2	40.0	6145	6 I77211	I77211 Sequence 1
25	3062.2	40.0	6145	6 AR533381	AR533381 Sequence
26	3062.2	40.0	6145	6 BD105951	BD105951 Animal mo
27	3062.2	40.0	6145	12 SYNMLPLN6	M63653 Moloney mur
28	3050.2	39.9	6620	6 AR302094	AR302094 Sequence
29	3007	39.3	6259	12 SYNRRV	M77239 Cloning vec
30	3006.8	39.3	6620	6 AX823828	AX823828 Sequence
31	3005.2	39.3	6620	12 SYNMLPLN3	M28247 Retroviral
32	2980.6	38.9	6690	12 AB086389	AB086389 Retrovira
33	2960.2	38.7	5901	12 AB086388	AB086388 Retrovira
34	2959.8	38.7	6149	12 SYNMLPLN2	M28246 Retroviral
35	2946.6	38.5	5883	12 AB041927	AB041927 Retrovira
36	2939	38.4	5904	12 AB086384	AB086384 Retrovira
37	2899.2	37.9	6141	6 BD138562	BD138562 Expressio
38	2860.4	37.4	6094	12 AF113968	AF113968 Cloning v
39	2842.4	37.1	6072	12 D88622	D88622 Biotronic
40	2835.2	37.0	6837	12 SYNMOV2	M64754 Moloney mur
41	2783.8	36.4	7797	6 E23355	E23355 Virus vecto
42	2703	35.3	5865	6 AR164477	AR164477 Sequence
43	2687.4	35.1	5874	6 BD138565	BD138565 Expressio
44	2687.4	35.1	5874	12 SYNMLPLN4	M28248 Retroviral
45	2675.4	35.0	6283	6 AX133941	AX133941 Sequence

## ALIGNMENTS

RESULT 1  
LOCUS AF132211 6501 bp DNA linear SYN 18-MAY-1999  
DEFINITION Cloning vector pFBneo, complete sequence.  
ACCESSION AF132211  
VERSION AF132211.1 GI:4838522  
KEYWORDS Cloning vector pFBneo  
ORGANISM Cloning vector pFBneo  
REFERENCE 1 (bases 1 to 6501)  
AUTHORS Other sequences; artificial sequences; vectors.  
TITLE Felts,K. and Grafsky,A.J.  
JOURNAL Direct Submission  
Submitted (01-MAR-1999) Technical Services, Stratagene, 11011 N.  
Torrey Pines Rd., La Jolla, CA 92037, USA.  
FEATURES  
source  
1..6501  
Location/Qualifiers  
/organism="Cloning vector pFBneo"  
/mol\_type="genomic DNA"  
/specific\_host="Escherichia coli"  
/db\_xref="taxon:93464"  
209..760  
/note="5'"  
760..2046  
/note="psi+; extended viral packaging signal"  
2057..2086  
/note="multiple cloning site"  
2092..2680  
/note="IRES"  
2678..3487  
/gene="neo"  
2678..3487  
/gene="neo"  
/function="neomycin resistance"  
/codon\_start=1  
/transl\_table=11  
/product="aminoglycoside 3'-phosphotransferase"  
/protein\_id="AAD31025.1"  
/db\_xref="GI:4838524"  
/translations="MGSFRIEODGLHAGSPAAWVERLFGVDWAQQTIGCSDAAVFRL  
SAGQRVFLVFKTDLSGNLQAEARLSWLTATGVCACAAVLVDVVTAGRWLLIGEV  
PGQDLSSHLAPAEKVSINADAMRRRLTLPATCFPDQAKHRIERARTRMEAGLVDD





QY 2345 AAAGTAGACGGCATCGCAGCTTGGATACACGCGCCCACTGAAGGCTGCCGACCCCGGG 2404  
Db |||||  
QY 1964 AAAGTAGACGGCATCGCAGCTTGGATACACGCGCCCACTGAAGGCTGCCGACCCCGGG 2023  
Db |||||  
QY 2405 GGTGACCATCCTCTAGACTGCCATGGATGGAGCTGTATCATCTCTCTTCTTGTGTAGCAA 2464  
Db |||||  
QY 2024 GGTGACCATCCTCTAGACTGCCATGGATGGAGCTGTATCATCTCTCTTCTTGTGTAGCAA 2082  
Db |||||  
QY 2465 CAGCTACAGGTGTCCACTCCGACATCCAGCTGACCCAGAGCCCAAGCAG----- 2513  
Db |||||  
QY 2083 CCGCGATCCGGTATTTTCCACCATAATTGCCGTCTTTTGGCAATGTGAGGGCCCGGAAC 2142  
Db |||||  
QY 2514 -----CCTGACGCCAGCGTGGTGA CAGAGTGACCATCACTGTAAAGCCAGTCAAGAT 2568  
Db |||||  
QY 2143 CTGGCCCTGTCTTCTGACGAGCATTCCTAGGGGTCTTTCCCTCTGCCCAAGGAATGC 2202  
Db |||||  
QY 2569 GTGGGTACTCTGTAGCTTGTGTACACGAGAACCCAGTAAAGGCTC-CAAGCTGCTCAT 2627  
Db |||||  
QY 2203 AAGTCTGTGTGAATGTGTGAAGAGCAGTTCCTCTGGAAGCTTCTTGAAGCAAAACAA 2262  
Db |||||  
QY 2628 CTACTGGACATCCACCCGGCACACTGGTGTGCCAAGCAGATTCAGCGGTAGCGGTAGCGG 2687  
Db |||||  
QY 2263 CGTCTGTAGACCCCTTTGAGGACGCGGAACCCCACTGGGACAGGTGCTCTGCG 2322  
Db |||||  
QY 2688 TACCGACTTACCTTACCATCAGCAGCCTCCAGCCAGAGGACATCGCCACTTACTACTG 2747  
Db |||||  
QY 2323 GCCAAAGGCCACGTGTATAAGATACACTGTCAAAAGGCGGCACAAACCCAGTGCACGTTG 2382  
Db |||||  
QY 2748 CCAGCAATATAGCTCTATCGGTGTTCCGCCAAGGACCAAGGTGGNAATCAAAACGAGG 2807  
Db |||||  
QY 2383 TGAGTTGGATAGTTGTGGAAAGAGTCAAAAGGCTCTCCTCAAGGTATTTCAACAAAGGGC 2442  
Db |||||  
QY 2808 TGGCTCAGGATCGGGTGGATCCGGCTCTGTGGTCTCAGGATCGAGGTCCAACTGTGTGA 2867  
Db |||||  
QY 2443 TGAAGATGCCAAGAGTACCCATTTGTATGGATCTGATCTGGGGCTCGGTGCACAT 2502  
Db |||||  
QY 2868 GAGCGGTGGAGGTGTGTCAACCTGGCCGGTCCCTGCGCTGTCTGTCTCC--GCATCT 2925  
Db |||||  
QY 2503 GCTTTACATGTGTTTGTAGTCGAGGTAAACACGCTAGGCCCCCGAACACACGCGGAGCT 2562  
Db |||||  
QY 2926 GGCTTCGATTTCAACACATATGTGATGATGGTGGTGAGACGACCTGGAAAGGTCTT 2985  
Db |||||  
QY 2563 GGTTTTCTTTTGAACAAACACGATGATATATGGAACAAACATTTATTTCTGAAGAAGACT 2622  
Db |||||  
QY 2986 GAGTGGATTGGAGAAATTCATCCAGATAGCAGTACGATTAACATATGC--GCCGTCTCTAA 3043  
Db |||||  
QY 2623 ---TGGACACCAACTTCTCTGCGCTCGATTTCTCCACCTGTGTGCCACGCGGTTCAT 2679  
Db |||||  
QY 3044 AGGATAGATTTACATATTCGCGAGACAAACGCAAGAACACATTTGTTCTGTCMAATGGACA 3103  
Db |||||  
QY 2680 GGGATCGTTTCGCATGATTGAACAGATGGATTGCACGAGGTCTCCGGCGCTTGGGT 2739  
Db |||||  
QY 3104 GCCTGAGACCCGAAGACACCGGGGTCTATTTTGTGCAAGCTTTTACTTGGGTTCOCCT 3163  
Db |||||  
QY 2740 GGAGAGGCTATTCGGCTATGACTGGGCACACAGACAAATCGGCTGCTCTGATGCGCGGT 2799  
Db |||||  
QY 3164 GGTTTGCTTATTTGGGCCAAGGACCCCGGTCAACGCTCTCCAGTGTCTTAAGCCCAACACGA 3223  
Db |||||  
QY 2800 GTTCCGGCTGTCAACGCGCAGGGGCGCCCGGTCTTTTGTCAAGACCCAGCTGTCCGGTGC 2859  
Db |||||  
QY 3224 CGCCAGCGCCGACCAACACACGCGGCCACCATCGCTGCGAGCCCTGTCCTGCTCCCTGC 3283  
Db |||||  
QY 2860 CTTGAATG-----AACTGAGACAGAGGACGCGCGGTATCTGTGCTGCCACAGAGGGCG 2915  
Db |||||  
QY 3284 GCCCAGAGCGGGCTCGGCACGCGCGGGGGCGCAGTGCAACAGAGGGGCTGCACTTCG 3343  
Db |||||  
QY 2916 TTCTTTCGCGAGCTGTGCTGACGCTGTGTCACTGAAGCGGAAGGACTGTGCTGTATTGG 2975  
Db |||||  
QY 3344 CCCTGGATCCCAACTCTGCTACTCTGTGGATGGAATCCTCTTCATCTATGGTGTCAATTC 3403  
Db |||||  
QY 2976 GCGAAGTGCCTGGGCGAGATCTCCTGTCTCATCTCACTTGTCTCTGCCGAGAAAGTATCCA 3035

QY 3404 TCACCTGCTTGTCTCTGAGAGTGAAGTTTCAGCAGGAGCGCAGAGCCCGCGGTACCAGC 3463  
Db |||||  
QY 3036 TCATGGCTGATGCAATCGCGCGCTGCATACGCTTGTATCCGGCTACTCTGCCCATTCGACC 3095  
Db |||||  
QY 3464 AGGCGCAGAACCAAGCTCTATAACGAGCTCAATCTAGGACGAAGAGAGGAGTACGATGTTT 3523  
Db |||||  
QY 3096 ACCAAGCGAAACATCGC-----ATCGAGCGACACGTACTCTCGATGGAAGCGGTC 3146  
Db |||||  
QY 3524 TGGCAAGAGACGTTGGCGGGGACCTCGATAGATGGGGGGGAAACCCGAGAAAGAACCCCTC 3583  
Db |||||  
QY 3147 TTGTGATCAGGATGATCTGGACGAAGAGCATCAGG--GCTCGCGCAGCGCAACTGT 3204  
Db |||||  
QY 3584 AGGAAGSCCTGTCAATGAATGACAGAAAGATAGATGGGGAGGCGCTTACAGTGAGATTG 3643  
Db |||||  
QY 3205 CGCCAGCTCAAGCGCGCATGCCGACGGGAGGATCTCGTCTGACCCATGCGCATGTC 3264  
Db |||||  
QY 3644 GGATGAAGGCGAGCGCGGAGGGGCAAG--GGACGATGGCTTTTACAGGCTCTCAG 3701  
Db |||||  
QY 3265 CTGCTTCCGATATCATGTGGGAAATGGCGCTTTTCTGGATTCATCGACTGTGGCG 3324  
Db |||||  
QY 3702 TACAGCCACCAAGGACACCTACGACGCCCTTTCACATGACGAGCCCTGCCCCCTCGTAA-- 3759  
Db |||||  
QY 3325 GCTGGGTGTGGCGACCGCTATCAGGACATAGCGTTGGCTACCGTGATATTGCTGAAGA 3384  
Db |||||  
QY 3760 --CTCGACGGCGCGCGGATCCGGATTCGAAATTTGTTAAAGACAGGATATCAGTGGT 3817  
Db |||||  
QY 3385 GCTTGGCGGAAATGGCTGACCGCTTCTCGTCTTTACGGTATCGCGCTCCGATTC 3444  
Db |||||  
QY 3818 CCAGGCTCTAGTTTGTACTCAACATATCACGACTGAAGCTATAGATAGGAGCCATA 3877  
Db |||||  
QY 3445 GCAGCGCATCGCTTCTATCGCTTCTTGACGAGTCTTCTGACGGGACTCTGGGGTTC 3504  
Db |||||  
QY 3878 GATAAATAAAGATTTTATTTAGTCTCCAGAAAAGGGGGGAATGAAAGACCCACCTG 3937  
Db |||||  
QY 3505 GATAAATAAAGATTTTATTTAGTCTCCAGAAAAGGGGGGAATGAAAGACCCACCTG 3564  
Db |||||  
QY 3938 TAGTTTGGCAAGCTAGCTTAAGTAAGCCATTTTTCAGAGGCTGG--AAAATACATACT 3996  
Db |||||  
QY 3565 TAGTTTGGCAAGCTAGCTTAAGTAAGCCATTTTTCAGAGGCTAGGAAAAATACATACT 3624  
Db |||||  
QY 3997 GAGNATAGAGATTCAGATCAAGTTAGNACAGA--GAGNACAGNATATGGCCAA 4055  
Db |||||  
QY 3625 GAGAATAGAGATTCAGATCAAGTTCAGGAAACAGATGGAACAGCTGAATATGGGCCAA 3684  
Db |||||  
QY 4056 CAGGATATCTGTGTAGCAGTTCCTGCCCC--GCTCAGGGCCAAAGAACAGTTGGAACAGG 4114  
Db |||||  
QY 3685 CAGGATATCTGTGTAGCAGTTCCTGCCCCGCTCAGGGCCAAAGAACAGATGGAACAGC 3744  
Db |||||  
QY 4115 AGAATATGGCCCAACAGGATATCTGTGTAAGCAGTTCCTGCCCCGCTCAGGGCCAAAG 4174  
Db |||||  
QY 3745 TGAATATGGCCCAACAGGATATCTGTGTAAGCAGTTCCTGCCCCGCTCAGGGCCAAAG 3804  
Db |||||  
QY 4175 AACAGATGTCCTCAGATGCGGTCCCGCTCAGCAGTTCCTGAGAACCATCAGATGTT 4234  
Db |||||  
QY 3805 AACAGATGTCCTCAGATGCGGTCCAGCTCAGCAGTTCCTGAGAACCATCAGATGTT 3864  
Db |||||  
QY 4235 TCCAGGGTGCCTCAGAGCCTGAAATGACCTGCTGCTTATTTTGAATCAATCAATCAGTT 4294  
Db |||||  
QY 3865 TCCAGGGTGCCTCAGAGCCTGAAATGACCTGCTGCTTATTTTGAATCAATCAATCAGTT 3924  
Db |||||  
QY 4295 CGCTTCTCGTCTTGTGTCGGCGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACAC 4354  
Db |||||  
QY 3925 CGCTTCTCGTCTTGTGTCGGCGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACAC 3984  
Db |||||  
QY 4355 CTTCACTCGCGCGCTCAGTCTCCGATAGATGCGGTCCCGCGGTACCCGTGTTCTCAAT 4414  
Db |||||  
QY 3985 CTTCACTCGCGCGCTCAGTCTCCGATGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4044  
Db |||||  
QY 4415 AAACCTCTTTCAGATTCATCCGATCGCTGCTGCTGCTTCTTGGAGGCTCT-CTCT 4473  
Db |||||  
QY 4045 AAACCTCTTTCAGATTCATCCGATCGCTGCTGCTGCTTCTTGGAGGCTCTCTCTCT 4104  
Db |||||  
QY 4474 GAGTGTGACTACCCCGTACGCGGGTCTTTTCACTTCTCCACCTACACAGGCTCTCACT 4533



```
Qy 6693 TGCATAATTCCTTACTGTCTATGCGCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACT 6752
Db 5933 TGCATAATTCCTTACTGTCTATGCGCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACT 5992
Qy 6753 CAACCAAGTCAATCTGAGAAATAGTATGCGCGCACCGAGTTGCTCTTTGCCCGCGTCACAA 6812
Db 5993 CAACCAAGTCAATCTGAGAAATAGTATGCGCGCACCGAGTTGCTCTTTGCCCGCGTCACAA 6052
Qy 6813 TACGGGATAATACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAAAAAGTT 6872
Db 6053 CACGGGATAATACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAAAAAGTT 6112
Qy 6873 CTTCCGGGCGGAAACTCTCAAGATCTTACCGCTGTTGAGATCGAGTTCGATGTAACCCA 6932
Db 6113 CTTCCGGGCGGAAACTCTCAAGATCTTACCGCTGTTGAGATCGAGTTCGATGTAACCCA 6172
Qy 6933 CTCTGTGACCCCACTGATCTTACGATCTTTTACTTTTACCCAGCGTTTCTGGGTGAGCAA 6992
Db 6173 CTCTGTGACCCCACTGATCTTACGATCTTTTACTTTTACCCAGCGTTTCTGGGTGAGCAA 6232
Qy 6993 AAACAGGAAGCAAAATGCCGCAAAAAGGGAATAAGGGCGACACGGAATGTTGAATAC 7052
Db 6233 AAACAGGAAGCAAAATGCCGCAAAAAGGGAATAAGGGCGACACGGAATGTTGAATAC 6292
Qy 7053 TCATACCTCTCTCTTTTCAATATATTGAAGCATTTATCAGGCTTATGCTCTCATGAGCG 7112
Db 6293 TCATACCTCTCTCTTTTCAATATATTGAAGCATTTATCAGGCTTATGCTCTCATGAGCG 6352
Qy 7113 GATACATATTTGAATGATTTAGAAAATAAACAAATAGGGGTTCGCGGCACATTTCCCC 7172
Db 6353 GATACATATTTGAATGATTTAGAAAATAAACAAATAGGGGTTCGCGGCACATTTCCCC 6412
Qy 7173 GAAAAGTGCCACCTGACGCTCTAAGAAACCAATATTATCATGACATTAACCTATAAAAAATA 7232
Db 6413 GAAAAGTGCCACCTGACGCTCTAAGAAACCAATATTATCATGACATTAACCTATAAAAAATA 6472
Qy 7233 GCGGTATCAGAGGCCCTTTGCTCT 7257
Db 6473 GCGGTATCAGAGGCCCTTTGCTCT 6497

RESULT 2
AX663075
LOCUS AX663075 9320 bp DNA linear PAT 24-MAR-2003
DEFINITION Sequence 20 from Patent WO02070740.
ACCESSION AX663075
VERSION AX663075.1 GI:29169369
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE
1. Wiesmueller L.
AUTHORS Test system for determining gene toxicities
TITLE Patent: WO 02070740-A 20 12-SEP-2002;
JOURNAL Wiesmueller, Lisa (DE)
FEATURES
source Location/Qualifiers
1..9320
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Plasmid p5-Puro-CMV- (N'-EGFP) -CMV-Red- (EGFP-EJ) "
misc_feature 1..1592
/note="Retroviral Vektor p5NM"
1617..2216
/note="unnamed protein product; Puromycin-Resistenzgen aus pRetroOn (Clontech, Palo Alto, CA, USA)"
CDS
1617..2216
/codon_start=1
/transl_table=1
/protein_id="CAD80174.1"
/db_xref="GI:29169370"
/translation="MTSEKPTVRLATRDDVPRAVRTLAAAFADYPATRTVDPDRHIE
```

## promoter

## CDS

```
RVTELQELFTRVGLDVGKVVADDCAAVAVWTTPTSEVAGAVFAETGPRMAELSGSR
LAQQQHEGLAHPKPAWFLATVGVSDHQKGLGSAVVLDLPGVEAAERAGVPAFL
ETSAENLPEYERLFTVTADVECPDRATWCMTRKPGA"
2267..2848
/note="CMV-Promotor aus pEGFP-N1 (Clontech, Palo Alto, CA,
USA)"
2906..>3349
/note="unnamed protein product; N'-EGFP, abgeleitet von
EGFP aus pEGFP-N1 (Clontech, Palo Alto, CA, USA)"
/codon_start=1
/transl_table=1
/protein_id="CAD80175.1"
/db_xref="GI:29169371"
/translation="MWSKGBELFTGVVPIILVELDGDVNGHKFVSVSGEGEDATYVKLT
LKPICTTGLKLPVPMPTLVTLTYGVQCFSRYPDHMKQHDFFKSGAMPEGYVQERTIFFK
DDGNYKTRAEVKEFGEDTLVNRIELKGIDFKEDGNILGHKLEYNVR"
3374..3392
/note="Retroviral Vektor p5NM"
3411..3992
/note="CMV-Promotor aus pDeRedi-N1 (Clontech, Palo Alto,
CA, USA)"
4038..4718
/note="unnamed protein product; Red aus pDeRedi-N1
(Clontech, Palo Alto, CA, USA)"
/codon_start=1
/transl_table=1
/protein_id="CAD80176.1"
/db_xref="GI:29169372"
/translation="MVRSSKNVKEPMRFKVRMEGTVNGHEFEIEGEGEGRPYRGHNT
VKLVKVGKGLPANDILSPQFOYGVSKVYVHPADIPDYKLSFPPEGFKVERVMNFD
GVVTVQDSLDQDGFYKVKFIGNPDSGPMVKTKTGWNEASTBRLTFRDQVLKG
EIHKALKDKDGGHYLVEFKSIYMAKPKVQLPGYYVDSKLDITSHNEDYITVEQYRT
EGRHHLFL"
```

## misc\_feature

## misc\_feature

## ORIGIN

```
Query Match 43.8%; Score 3349; DB 6; Length 9320;
Best Local Similarity 94.4%; Pred.No. 0;
Matches 3570; Conservative 0; Mismatches 150; Indels 61; Gaps 7;
Qy 3772 GCGGATCCGGATAGTCCAAATTTGTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTT 3831
Db 5541 GGGCCCGCGATTAGTCCAAATTTGTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTT 5600
Qy 3832 TGACTCAACAATATCACCAGCTGAAGCCTATAGAGTACGAGCCATAGATAAATAAAGA 3891
Db 5601 TGACTCAACAATATCACCAGCTGAAGCCTATAGAGTACGAGCCATAGATAAATAAAGA 5660
Qy 3892 TTTTATTAGTCTCCAGAAAAGGGGGAATGAAGACCCCACTGTAGTTTCCGCAAGC 3951
Db 5661 TTTTATTAGTCTCCAGAAAAGGGGGAATGAAGACCCCACTGTAGTTTCCGCAAGC 5720
Qy 3952 TAGCTTAAGTAAGCCCATTTTTCGAAGGCATCGAAAATACATACTGAGAATAGAGAAGTT 4011
Db 5721 TAGCTTAAGTAAGCCCATTTTTCGAAGGCATCGAAAATACATACTGAGAATAGAGAAGTT 5780
Qy 4012 CAGATCAAGTTTAGGAACACAGAGACAGAGATATATGGGCCAAACAGAGATATCTGTGTA 4071
Db 5781 CAGATCAAGTTTAGGAACACAGAGACAGAGATATATGGGCCAAACAGAGATATCTGTGTA 5840
Qy 4072 AGCAGTTCCTGCCCC -GCTCAGGGCCAGAACACAGTTGGACACAGAGATATGGGCCAAC 4130
Db 5841 AGCAGTTCCTGCCCCCGGCTCAGGGCCAGAACACAGTTGGACACAGAGATATGGGCCAAC 5900
Qy 4131 AGGATATCTGTGTAGCAGTCTCTGCCCGGCTCAGGGCCAGAACACAGATGTGTCCCGAG 4190
Db 5901 AGGATATCTGTGTAGCAGTCTCTGCCCGGCTCAGGGCCAGAACACAGATGTGTCCCGAG 5960
Qy 4191 ATGGGTCCTCCGCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCCCAAG 4250
```

Db	5961	ATGCGTCCCGCCTCAGCAGTTTCTAGAAACCATCAGATGTTTCCAGGGTGCCCAAG	6020
Qy	4251	GACCTGAATGACCCCTGTGCTTATTTGAACCAATCAGTTCGCTTCTGCTCTGT	4310
Db	6021	GACCTGAATGACCCCTGTGCTTATTTGAACCAATCAGTTCGCTTCTGCTCTGT	6080
Qy	4311	TGCGGGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGCGGC	4370
Db	6081	TGCGGGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGCGGC	6140
Qy	4371	AGTCTCTCGATGACTGCGTCCCGGGTACCGGTGTTCTCAATAAACCCCTCTTGCAATT	4430
Db	6141	AGTCTCTCGATGACTGCGTCCCGGGTACCGGTGTTCTCAATAAACCCCTCTTGCAATT	6200
Qy	4431	GAATCCGACTGCTGCTGCTGCTTCTCTGCGGGGTCT-CTCTGAGTGATGACTACCC	4489
Db	6201	GAATCCGACTGCTGCTGCTGCTTCTCTGCGGGGTCTCTCTGAGTGATGACTACCC	6260
Qy	4490	GTGAGC-GGGGTCTTTCAGTTCCTCCACCTACACAGGCTCACTAACATTCCTGATGG	4548
Db	6261	GTGAGCGGGGTCTTTCAGTTCCTCCACCTACACAGGCTCACTAACATTCCTGATGG	6320
Qy	4549	CGCAGGGACTCGCTCAGCCCGGTTTTTGTATATAAAATGCAAGACAGTGTCCCT	4608
Db	6321	CGCAGGGACTCGCTCAGCCCGGTTTTTGTATATAAAATGCAAGACAGTGTCCCT	6380
Qy	4609	TCAGCCAGACTACATCTCGACTCTCGGCTTTATAAAGAAATGTTGAAGGCTCTGTGA	4668
Db	6381	TCAGCCAGACTACATCTCGACTCTCGGCTTTATAAAGAAATGTTGAAGGCTCTGTGA	6440
Qy	4669	CTATCTGCCACAGGACTTTTAAGATTTTATGCTCTCCGATGAGGGATTTAGTCAATC	4728
Db	6441	CTATCTGCCACAGGAC-TTTTAAAGATTTTATGCTCTCCGATGAGGGATTTAGTCAATC	6499
Qy	4729	TATCTCTGCTATTTTGTGCTTCTCCGTATTTTAAATTTCTAGTTTGCATCCCTTCC	4788
Db	6500	TATCTCTGCTATTTTGTGCTTCTCCGTATTTTAAATTTCTAGTTTGCATCCCTTCC	6559
Qy	4789	TGAGACACGGGATTTGACAGTAGTTAATATCTGAGGGCAGGCTTCTGTAAGGTT	4848
Db	6560	TGAGACACGGGATTTGACAGTAGTTAATATCTGAGGGCAGGCTTCTGTAAGGTT	6619
Qy	4849	GCTGGGCTCAGTGTGAGATTTTGCATATAAGGGTCTCCGCTCTGTTACAGACAG	4908
Db	6620	GCTGGGCTCAGTGTGAGATTTTGCATATAAGGGTCTCCGCTCTGTTACAGACAG	6679
Qy	4909	TCGGAACTAGAGTGATATCTCAGAGTCCCGCGGTTCCGGGGCTCTGATCTCAGGGCAT	4968
Db	6680	TCGGAACTAGAGTGATATCTCAGAGTCCCGCGGTTCCGGGGCTCTGATCTCAGGGCAT	6739
Qy	4969	CTTTGGCTAGAGATCTCTACGGCGAGGATGCTGGCGGGTACCGAGCTCGAATTCGT	5028
Db	6740	CTTTGGCTAGAGATCTCTACGGCGAGGATGCTGGCGGGTACCGAGCTCGAATTCGT	6799
Qy	5029	AATCATGGTCATAGCTGTTTC-----CTGTGTGAAATGTTTATCGCTCACA	5075
Db	6800	GCGGTTGCTGGCGCTTATATGCGCGACATCACCGATGGGGAGATCGGGCTCGCCACTTC	6859
Qy	5076	ATTCCACACAAATACGAGCCGGAAGCATAAAGTGTAAAGCTTGGGGTCCCTAATGAGT-	5134
Db	6860	GGGCTCATGAGCGCTTGTTCGGCGTGGGTATGGTGGCAGGCCCGTGGCGGGGACTG	6919
Qy	5135	-----GAGCTAACTCATTAAT	5152
Db	6920	TTGGGGCCCATCTCTTGATGACCAATTCCTTTGCGCGCGCGGTCTCAACGCGCTCAAC	6979
Qy	5153	TGCGTTGCTCAGTCCCGCTTTCAGTCCGGAACCTGTGTCGACGCTCATTAATG	5212
Db	6980	CTACTACTGGGCTGCTTCTTAATGACAGGATGTCATAAAGGAGAGCGTCTGCAITTAATG	7039
Qy	5213	AATCGGCCAAACGCGGGGAGAGCGGTTTTGCGTATTTGGCGCTCTTTCGCGCTTCTCGCT	5272
Db			
Db	7040	AATCGGCCAAACGCGGGGAGAGCGGTTTTGCGTATTTGGCGCTCTTTCGCTTCTCGCT	7099
Qy	5273	CACTCACTCGCTCGCTCGCTCGTTTGGTGGCGAGCGGTATCAGCTCACTCAAAAGC	5332
Db	7100	CACTCACTCGCTCGCTCGCTCGTTTGGTGGCGAGCGGTATCAGCTCACTCAAAAGC	7159
Qy	5333	GCTAATACGGTTATCCACAGAAATCAGGGGATAACCGAGSAAAGACATGTGAGCAAAAG	5392
Db	7160	GCTAATACGGTTATCCACAGAAATCAGGGGATAACCGAGSAAAGACATGTGAGCAAAAG	7219
Qy	5393	CCAGCAAAAGGCGCAGAAACCGTAAAGCCGCTTGTGCGGTTTTTTCATAGGCTCG	5452
Db	7220	CCAGCAAAAGGCGCAGAAACCGTAAAGCCGCTTGTGCGGTTTTTTCATAGGCTCG	7279
Qy	5453	CCCCCTGACGAGCATCAAAAATCGAGCTCAAGTCAAGGTGGCGGAAACCCGACAG	5512
Db	7280	CCCCCTGACGAGCATCAAAAATCGAGCTCAAGTCAAGGTGGCGGAAACCCGACAG	7339
Qy	5513	ACTATAAAGATACCAAGGTTTTCCCTCGAAAGCTCCCTCGTGGCTCTCTCTGTTCCGAC	5572
Db	7340	ACTATAAAGATACCAAGGTTTTCCCTCGAAAGCTCCCTCGTGGCTCTCTCTGTTCCGAC	7399
Qy	5573	CCTGCGCTTACCGATACCTGTCGCTTCTCTCTCGGAAAGGTGGCGGCTTCTCA	5632
Db	7400	CCTGCGCTTACCGATACCTGTCGCTTCTCTCTCGGAAAGGTGGCGGCTTCTCA	7459
Qy	5633	TAGCTCAGCTGTAGGTATCTCAGTTCCGTTAGGTGCTTCCCTCAAGCTGGGCTGT	5692
Db	7460	ATGCTCAGCTGTAGGTATCTCAGTTCCGTTAGGTGCTTCCCTCAAGCTGGGCTGT	7519
Qy	5693	GCAAGAAACCCCGTTACGCCCGAGCGCTTCCCTTATCCGTTAACTATCGTCTTGAGTC	5752
Db	7520	GCAAGAAACCCCGTTACGCCCGAGCGCTTCCCTTATCCGTTAACTATCGTCTTGAGTC	7579
Qy	5753	CAACCGGTAAGACACGACTTATGCCACTGCGCAGAGCACTGGTAAACAGATTAGCAG	5812
Db	7580	CAACCGGTAAGACACGACTTATGCCACTGCGCAGAGCACTGGTAAACAGATTAGCAG	7639
Qy	5813	ACGAGGTATGTAGCGGTGCTACAGATTTCTTGAAGTGGTGGCTTAACTACGGCTACAC	5872
Db	7640	ACGAGGTATGTAGCGGTGCTACAGATTTCTTGAAGTGGTGGCTTAACTACGGCTACAC	7699
Qy	5873	TAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAGT	5932
Db	7700	TAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAGT	7759
Qy	5933	TGGTAGCTTCTGATCCGCAAAACCAACCGCTGGTAGCGGTGTTTTTTTGGCAA	5992
Db	7760	TGGTAGCTTCTGATCCGCAAAACCAACCGCTGGTAGCGGTGTTTTTTTGGCAA	7819
Qy	5993	GCAGCAGATTACGCGCAGAAAAGGATCTCAAGAGATCCTTTGATCTTTTCTACGG	6052
Db	7820	GCAGCAGATTACGCGCAGAAAAGGATCTCAAGAGATCCTTTGATCTTTTCTACGG	7879
Qy	6053	GTCTCAGCTCAGTGGAAACGAACTCAGCTTAAAGGATTTTGGTTCATGAGATTACAA	6112
Db	7880	GTCTCAGCTCAGTGGAAACGAACTCAGCTTAAAGGATTTTGGTTCATGAGATTACAA	7939
Qy	6113	AAGGATCTTACCTAGATCTTTTAAATTTAAAGTAAATCAATCTAAAGTAT	6172
Db	7940	AAGGATCTTACCTAGATCTTTTAAATTTAAAGTAAATCAATCTAAAGTAT	7999
Qy	6173	ATATGAGTAACTTGGTCTGACAGTTACCAATGCTTAACTAGGACACCTATCTCAG	6232
Db	8000	ATATGAGTAACTTGGTCTGACAGTTACCAATGCTTAACTAGGACACCTATCTCAG	8059
Qy	6233	GATCTGTCTATTTCTGTTTCAATCATAGTTGCTGACTCCCGCTGCTGTAGATACTACGAT	6292
Db	8060	GATCTGTCTATTTCTGTTTCAATCATAGTTGCTGACTCCCGCTGCTGTAGATACTACGAT	8119
Qy	6293	ACGGAGGGCTTACCATCTGCGCCAGTGTCTGCAATGATACCGCAGAGACCCAGCTCACC	6352
Db	8120	ACGGAGGGCTTACCATCTGCGCCAGTGTCTGCAATGATACCGCAGAGACCCAGCTCACC	8179

```
Qy 6353 GCCTCCAGATTATCAGCAATAAACCCAGCAGCCGGAAGGGCCGAGCGCAGAGTGGTCC 6412
Db 8180 GCCTCCAGATTATCAGCAATAAACCCAGCAGCCGGAAGGGCCGAGCGCAGAGTGGTCC 8239
Qy 6413 TGCAACTTTATCCCGCTCCATCCAGTCTATTAAATTTGTCGGGAAGCTAGAGTAAGTAG 6472
Db 8240 TGCAACTTTATCCCGCTCCATCCAGTCTATTAAATTTGTCGGGAAGCTAGAGTAAGTAG 8299
Qy 6473 TTCGCCAGTTAATAGTTTTCGCAACGTTGTTGCCATTGCTACAGGC-TCGTGGTGTACG 6531
Db 8300 TTCGCCAGTTAATAGTTTTCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACG 8359
Qy 6532 CTCGTGTTTGGTATGCTTCAATCAGCTCCGGTTCCCAAGATCAAGGCGAGTTACATG 6591
Db 8360 CTCGTGTTTGGTATGCTTCAATCAGCTCCGGTTCCCAAGATCAAGGCGAGTTACATG 8419
Qy 6592 ATCCCCCATGTTGTGCAAAAAGCGTTAGTCTTCCTCGGTCTCCGATCGTTGTCAGAG 6651
Db 8420 ATCCCCCATGTTGTGCAAAAAGCGTTAGTCTTCCTCGGTCTCCGATCGTTGTCAGAG 8479
Qy 6652 TAAGTTGCCCGCAGTGTATCACTCATGTTATGCGACACTGCAATAATCTCTTACTGT 6711
Db 8480 TAAGTTGCCCGCAGTGTATCACTCATGTTATGCGACACTGCAATAATCTCTTACTGT 8539
Qy 6712 CATGCCATCCGTAAGATGCTTTCTGTGACTGGTGAGTACTCAACCAAGTCAATCTGAGA 6771
Db 8540 CATGCCATCCGTAAGATGCTTTCTGTGACTGGTGAGTACTCAACCAAGTCAATCTGAGA 8599
Qy 6772 ATAGTGATGCGGACCGAGTTGCTCTGTCGCCGCGTCAATACGGGATATACCGGCC 6831
Db 8600 ATAGTGATGCGGACCGAGTTGCTCTGTCGCCGCGTCAATACGGGATATACCGGCC 8659
Qy 6832 ACATAGCAGAACTTTAAAGTGCTCATCATTTGGAAGAGTTCTTCGGGGCGAAACTCTC 6891
Db 8660 ACATAGCAGAACTTTAAAGTGCTCATCATTTGGAAGAGTTCTTCGGGGCGAAACTCTC 8719
Qy 6892 AAGGATCTACCGTGTGAGATCAGTTCGATGTAACCCACTCGTGACCCAACTGATC 6951
Db 8720 AAGGATCTACCGTGTGAGATCAGTTCGATGTAACCCACTCGTGACCCAACTGATC 8779
Qy 6952 TTCAGCATCTTTACTTTTACAGGTTTCCGCGACATTTCCGGTGAGCAAAAACAGGAGGCAAAATGC 7011
Db 8780 TTCAGCATCTTTACTTTTACAGGTTTCCGCGACATTTCCGGTGAGCAAAAACAGGAGGCAAAATGC 8839
Qy 7012 CGCAAAAAGGGAATAAGGCGACACGAAATGTTGAATATCTCATCTCTCTTTTCA 7071
Db 8840 CGCAAAAAGGGAATAAGGCGACACGAAATGTTGAATATCTCATCTCTCTTTTCA 8899
Qy 7072 ATATTATTGAAGCATTTATCAGGTTATTGTCTCATGACGCGATACATATTGTAATGTAT 7131
Db 8900 ATATTATTGAAGCATTTATCAGGTTATTGTCTCATGACGCGATACATATTGTAATGTAT 8959
Qy 7132 TTAGAAAATAAACAAATAGGGTTCCGCGACATTTCCCGGAAAGTGCCACCTGACGT 7191
Db 8960 TTAGAAAATAAACAAATAGGGTTCCGCGACATTTCCCGGAAAGTGCCACCTGACGT 9019
Qy 7192 CTAAGAAACCAATTATTATCATGACATTAACCTATATAAATAAGGGTATCAGAGGCCCTT 7251
Db 9020 CTAAGAAACCAATTATTATCATGACATTAACCTATATAAATAAGGGTATCAGAGGCCCTT 9079
Qy 7252 TCGTCTCGCGGTTTCGGTGATGACGGTGAAAACTCTGACACATGACGTCCCGGAGAC 7311
Db 9080 TCGTCTCGCGGTTTCGGTGATGACGGTGAAAACTCTGACACATGACGTCCCGGAGAC 9139
Qy 7312 GGTACAGCTGTCTGTAAGCGATGCGGAGCAGACAACCGCTCAGGCGCGGTGACG 7371
Db 9140 GGTACAGCTGTCTGTAAGCGATGCGGAGCAGACAACCGCTCAGGCGCGGTGACG 9199
Qy 7372 GGGTGTGGCGGGTGTGCGGGCTGGCTTAACCTATGCGGCATCAGAGCAGATTGACTGAG 7431
Db 9200 GGGTGTGGCGGGTGTGCGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGACTGAG 9259
```

```
Qy 7432 AGTCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAATACCGCATCAG 7491
Db 9260 AGTCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAATACCGCATCAG 9319
Qy 7492 G 7492
Db 9320 G 9320

RESULT 3
LOCUS E23356 6644 bp DNA linear PAT 18-JUN-2001
DEFINITION Virus vector system expressing apoptosis-related gene.
ACCESSION E23356
VERSION E23356.1 GI:13024379
KEYWORDS JP 199075859-A/2.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 6644)
AUTHORS Hirofumi,H.
TITLE Virus vector system expressing apoptosis-related gene
JOURNAL Patent: JP 199075859-A 2 23-MAR-1999;
R B R JENSERU KK
COMMENT OS Unidentified
PN JP 199075859-A/2
PD 23-MAR-1999
PF 08-SEP-1997 JP 1997259235
PR HIROFUMI HAMADA
PI C12N15/09,C12N5/10,C12N7/00//A61K35/76,A61K48/00,(C12N5/10,PC
C12R1:91),
PC (C12N7/00,C12R1:92),C12N15/00,C12N5/00,(C12N5/00,C12R1:91) CC
Strandedness: Double;
CC Topology: Linear;
FH Key 1..6644 Location/Qualifiers
FT source 1..6644 /organism='Unidentified'.
FEATURES
source
Location/Qualifiers
1..6644
/organism='unidentified'
/mol_type='genomic DNA'
/db_xref='taxon:32644'

ORIGIN
Query Match 43.2%; Score 3307.8; DB 6; Length 6644;
Best Local Similarity 72.9%; Pred. No. 0;
Matches 4739; Conservative 0; Mismatches 1352; Indels 410; Gaps 21;
Qy 843 TTCTAGAGAACCATCATCAGATGTTCCAGGGTCCGCCAAGGACCTGAAATGACCCCTGTGCT 902
Db 467 TGCTAGAGAACCATCATCAGATGTTCCAGGGTCCGCCAAGGACCTGAAATGACCCCTGTGCT 526
Qy 903 TATTGAACTAACCAATCAGTTCGCTTCTGCTTCTGTCGCGCTTCTGCTCCCGAG 962
Db 527 TATTGAACTAACCAATCAGTTCGCTTCTGCTTCTGTCGCGCTTCTGCTCCCGAG 586
Qy 963 CTCAATAAAGAGCCCAACACCCCTCACTCGCGCGGCGAGTCTCCGATAGACTGCGTCC 1022
Db 587 CTCAATAAAGAGCCCAACACCCCTCACTCGCGCGGCGAGTCTCCGATAGACTGAGTCC 646
Qy 1023 CCGGGTACCGGTATTCCCAATAAGCTCTTGTGTTTGTGATCCGAATCGTGACTCGC 1082
Db 647 CCGGGTACCGGTATTCCCAATAAGCTCTTGTGATCCGAATCGTGACTCGC 705
Qy 1083 TGATCTTGGAGGGTCTCTCAGATTGATGACTGCC-ACCTCGGGGTCTTTCA TTT 1141
Db 706 TGTTCTTGGAGGGTCTCTCTCTGAGTGATTGACTACCCGTCAGCGGGTCTTTCA TTT 765
Qy 1142 GGAGTTCCACCGAGATTGGAGACCCCTGCGCAGGAGCACCGACCCCCCGCGGGAG 1201
Db 766 GGGGGTCTGTCGCGGATCGGAGACCCCTGCGCAGGAGCACCGA-CCCACACCGGGAG 824
```



Qy	1202	GTAAAGCTGGCCAGCAACTTATCTGTGTCGTCCGATTGTCTAGTGTCTATGACTGATTTT	12661
Db	825	GTAAAGCTGGCCAGCAACTTATCTGTGTCGTCCGATTGTCTAGTGTCTATGACTGATTTT	884
Qy	1262	ATCGGCTCGGTGGGTACTAGTTAGCTTAACTAGCTCTGTATCTGGCGGACCCGTGCTGGA	1321
Db	885	ATCGGCTCGGTGGGTACTAGTTAGCTTAACTAGCTCTGTATCTGGCGGACCCGTGCTGGA	944
Qy	1322	ACTGACGAGTTTCGGAAACACCCGGCCGCAACCCCTGGGAGACGTCCTCCAGGGAGACTTCGGGGGC	1381
Db	945	ACTGACGAGTTTCGGAAACACCCGGCCGCAACCCCTGGGAGACGTCCTCCAGGGAGACTTCGGGGGC	1004
Qy	1382	CGTTTTTGTGGCCCGACCTGAGTCTCTAAATATCCCGATCGTTTTAGGACTCTTTGGTGCACC	1441
Db	1005	CGTTTTTGTGGCCCGACCTGAGTCTCTAAATATCCCGATCGTTTTAGGACTCTTTGGTGCACC	1064
Qy	1442	CCCCCTAGAGGGGATATGTCTCTGGTAGGAGCGAGAACTTAACACAGTTCCCGCC	1501
Db	1065	CCCCCTAATAGAGGGGATATGTGTCTCTGGTAGGAGCGAGAACTTAACACAGTTCCCGCC	1122
Qy	1502	TCGGTCTGAATTTTGTCTTTCGGTTTGGGACCGAAGCCGCGCCGCGCTCTTCTCTGCTG	1561
Db	1125	TCGGTCTGAATTTTGTCTTTCGGTTTGGGACCGAAGCCGCGCGCTCTTCTCTGCTG	1184
Qy	1562	CAGCATCGTTCTGTGTTGTCCTGTCTGCTGACCTGTGTTCTGTATTTGTCTGAAATATGG	1621
Db	1185	CAGCATCGTTCTGTGTTGTCCTGTCTGCTGACCTGTGTTCTGTATTTGTCTGAAATATGG	1244
Qy	1622	CCGGGGCTAGACTGTACACTCGCTTTAAGTTTGAACCTTAGGTCTACTGGAAAGATCTCGA	1681
Db	1245	CC-----AGACTGTATTACCACTCCCTTAAAGTTTGAACCTTAGGTCTACTGGAAAGATCTCGA	1298
Qy	1682	GGCGATCGCTCAACAACAGTCGGTAGATGTCAAGAGAGAGACTTGGTTTACCTTCTGCTC	1741
Db	1299	GGCGATCGCTCAACAACAGTCGGTAGATGTCAAGAGAGAGACTTGGTTTACCTTCTGCTC	1358
Qy	1742	TGCAGAAATGGCCAACTTTAAAGCTCGGATGGCCGCGAGACGGACCTTTAAACCGAGACCT	1801
Db	1359	TGCAGAAATGGCCAACTTTAAAGCTCGGATGGCCGCGAGACGGACCTTTAAACCGAGACCT	1418
Qy	1802	CATCACCCAGGTTAAGATCAAGGTCCTTTTCACTGGCCCGCATGGACACCCAGACAGGT	1861
Db	1419	CATCACCCAGGTTAAGATCAAGGTCCTTTTCACTGGCCCGCATGGACACCCAGACAGGT	1478
Qy	1862	CCCTCATCATCGTGACCTGGGAAGCCTTGGCTTTTGTGACCCCTTCCCTGGGTCAAGCCCTT	1921
Db	1479	CCCTCATCATCGTGACCTGGGAAGCCTTGGCTTTTGTGACCCCTTCCCTGGGTCAAGCCCTT	1538
Qy	1922	TGTACACCTTAAGCCTCCGCTCTCTTCTCTCATCGCCCGCTCTCTCCCCCTTGAACC	1981
Db	1539	TGTACACCTTAAGCCTCCGCTCTCTTCTCTCATCGCCCGCTCTCTCCCCCTTGAACC	1598
Qy	1982	TCCTCGTTTGACCCCGGCTCGATCCTCCCTTTATTCAGGCCCTCACTCCTTCTCTAGGCGC	2041
Db	1599	TCCTCGTTTGACCCCGGCTCGATCCTCCCTTTATTCAGGCCCTCACTCCTTCTCTAGGCGC	1658
Qy	2042	CCCCATATGGCCATATCAGATCTTATATGGGACACCCCGCCCTTGTAACCTTCCTCTGA	2101
Db	1659	-----CCATATGAGATCTTATATGGGACACCCCGCCCTTGTAACCTTCCTCTGA	1708
Qy	2102	CCCTGACATGACAAGAGTTACTAACAGCCCTCTCTCTCAAGCTCACTTACAGCT--TCTA	2160
Db	1709	CCCTGACATGACAAGAGTTACTAACAGCCCTCTCTCTCAAGCTCACTTACAGGCTCTCTA	1768
Qy	2161	CTTAGTCCAGCAAGAGTCTGGAGACTCTGGGGGAGCCCTACCAAGAAACAATCTGACCG	2220
Db	1769	CTTAGTCCAGCAAGAGTCTGGAGACTCTGGGGGAGCCCTACCAAGAAACAATCTGACCG	1828
Qy	2221	ACCGGTGGTACTTCAACCTTACCGAGTCGGGACACAGTGTGGGTCCGCGACACACAGAC	2280
Db	1829	ACCGGTGGTACTTCAACCTTACCGAGTCGGGACACAGTGTGGGTCCGCGACACACAGAC	1888
Qy	2281	TAAGAACTCTAGAACTCGCTGGAAAGAGACCTTACACAGTCTCTGTGTACCAACCCCAACCGC	2340

Db	1989	TAAGAACTTAGAACCTCGCTGGAAAGACCTTACACAGTCTGCTGACCAACCCCAACCCG	1948
Qy	2341	CCTCAAAAGTAGACGGCATCGCAGCTTGGATACACGCCGCCACAGCTGAAGGCTGCCACCC	2400
Db	1949	CCTCAAAAGTAGACGGCATCGCAGCTTGGATACACGCCGCCACAGCTGAAGGCTGCCACCC	2008
Qy	2401	CGGGGTGGACCAATCCTCTAGACTGCCATCGGBATGGAGCTGTATCATCTCTTCTTGGTA	2460
Db	2009	CGGGGTGGACCAATCCTCTAGACTGCCCGGATCCAGTGTGGTGGTAGGGAATTCGAGCTC	2068
Qy	2461	GCAACAGCTACAGGTGTCCACTCCGACATCCAGCTGACCCAGAGCCCAAGCAGCCTGAGC	2520
Db	2069	GGTACCCGGGGCGCGCGGATCCTTAAATTAAAGCGCCGCGTAACAA	2116
Qy	2521	GCCAGCGTGGGTGACAGAGTGACCATCACCTGTAAAGGCCAGTCAGGATGTGGGTACTTCT	2580
Db	2117	-----TTG	2119
Qy	2581	GTAGCTTTGGTACCAGCAGAAAGCCAGGTAAAGCTCCAAAGCTGTGTATCTACTTGGACATCC	2640
Db	2120	TTAACTAACTAACTAAAGCTAGCAACGGTTTCCCTCTAGCGGGATCAATTCCTCCCTCTCCC	2179
Qy	2641	ACCCGGCACACTGGTGTGCCAAGCAGATTCAGCGGTAGCGGTAGCGGTACCGACTTCACC	2700
Db	2180	TCCCCCCCCCCCTAACGTTACTTGGCCGA	2206
Qy	2701	TTACCATCAGCAGCCTCCAGCCAGAGGACATCGCCACTACTACTGCCAGCAATATAGC	2760
Db	2207	-----AGCCGCTTGGAAATAGGCCGGTGTGGTTGCTATATGTTATTTTCACC	2257
Qy	2761	CTCTATCGGTCTGTTTCGGCCAAAGGACCAAGGTGGAAATCAAAAGAGTGGCTCAGGATCG	2820
Db	2258	ATAATTGCCGCTTTTGGCAATGTGA	2282
Qy	2821	GGTGGATCCGGCTCTGGTGGCTCAGATCGAGGTCCAACTGTGGAGAGCGGTGGAGGT	2880
Db	2283	-----GGGCCCGGAAACCTG	2297
Qy	2881	GTTGTGCAACTGGCCGGTCCCTGGCGCTGTCCTGCTCCGATCTGGCTTCGATTTCACC	2940
Db	2298	GCCCTGTCTCTTGACAGACATCTCTAGGGCTCTTCCCTCTCCGCAAGGATGCAAG	2357
Qy	2941	ACATATTGGATGAGTTGGGTGAGACAGGCACCTGGAAAGGTCTTGAGTGGATTGGAGAA	3000
Db	2358	GTCTGTTGAATGCTGTGAAGGAGCAGTCTCTCTGGAAGTCTCTTGAAGACAAACAACGT	2417
Qy	3001	ATTCAATCCAGATAGCAGTACGATTAACTATGCGCGTCTCTAAAGGATAGATTTACAATA	3060
Db	2418	CTGTAGGCA---CCCTTTGAGGAGCAGGGAACCCGCCACCTGGCGACAGGTGCCTCTGCG	2474
Qy	3061	TCGCGAGACCAAGCCAAAGAACACATTGTTCTGCAATGGACAGCCTGAGACCCGAAAGAC	3120
Db	2475	GCCAAAAGCACAGTGTATAGATACACTGTGAAGGGCGGACAAACCCAGTGCCACGTTG	2534
Qy	3121	ACCGGGGTCTATTTTGTGCAAGCTTTACTTCGGCTTCCCTCGGTTTGGTTTATTTGGGGC	3180
Db	2535	TGAGTTGGATGTTGTGMAAGAGTCAATGGCTCTCCTCAAGCGTATTCAAACAAGGGGC	2594
Qy	3181	CAAGGACCCCGGTCAACGGTCTCCAGTGTCTAAGCCCAACGAGCGCCAGCGCCGACCA	3240
Db	2595	TGAAGGATGCCAGAGGTACCCCATTT-----	2621
Qy	3241	CCAACACGGGGCCCAACCATCGCGTCGAGCCCTGTCTCCCTGGCCAGGCGGCTCGG	3300
Db	2622	-----GTATGGG	2628
Qy	3301	CCAGCGCGGGGGCGCAGTGACACAGAGGGGTGGAATTCGCCCTGGATCCCAAACTC	3360
Db	2629	ATCTGATCTGGGGCTCGGTGCACATGCTTTACATGTTTGTAGTCGAGGTTAAAAAAGT	2688
Qy	3361	TGCTACTGTGTGGANGAAATCTCTTTCATCTATAGGTGTCAATTCACACTGCTTGTCTCTG	3420



Db 2689 CTAGGCCCCCGAA-----CCACGGGACGTGGTTTCTCTTTGAAAAACACGATATA 2741  
Qy 3421 AGAGTGAAGTTTCAGCAGAGCGCAGAGCCCGCCGCTACACAGAGGGCCAGAACGAGCTC 3480  
Db 2742 CCATGGTTCATGAAACATATTTAACATTTCTCAACAGATCTAGAATTAGTAGAGTAGCGA 2801  
Qy 3481 TATAACGAGCTCAATCTAGGACGAGAGAGAGAGTACGATGTTTTTGGACAGAGACGTGGC 3540  
Db 2802 CAGAGAAGATTACATGCTTTTATGAGGATAATAACATCATGTGGAGCGGCAATTTCGTA 2861  
Qy 3541 CGGACCCCTGAGATGGGGGAAAGCCGAGAGAGGAAGAACCCCTCAGAGAGGCTCTGACAA 3600  
Db 2862 CGAATAACAGGAGAAATCATTTCCGACGTACATATTTGAAGCTATATAGGACGATAGCTG 2921  
Qy 3601 GAACCTGCAGAAAGATAGATGGCGGAGCCCTACAGTGAGATGTTGGATGAAAGCGAGCGC 3660  
Db 2922 TTTGTGCAGAGCCATTCGATTTGGTAG--TGCAGTTTTCGAATGGAACAAAGGATTTTGA 2979  
Qy 3661 CGGAGGGCAGGGCCAGATGGCTTTTACCAGGGTCTCAGTACAGCCACCAGGACACC 3720  
Db 2980 CACGATTGTAGCTGTAGACACCCCTTATTTGACGAAGTAGATAGAAATTCGAGTGGT 3039  
Qy 3721 TACGACGCCCTTCACATCAGCGCCCTGCCCTCGCTAACTCGACGGCGCGCGATCCG 3780  
Db 3040 AAGTCCTTGTGTATGTAGG-----GAGTTGATTTCAGACTATGCACCA 3085  
Qy 3781 GATTAGTCCAAATTTGTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTGTACTCAAC 3840  
Db 3086 GATTGTTTGTGTTAATAGAAATGAATGGCAAGTTAGTCAAACTACGATTGAAGAACTC 3145  
Qy 3841 AATATCACCAGCTGAGCTATAGAGTACGAGCAGCATAGATAAATAAAGATTTTATTTA 3900  
Db 3146 ATTCCACTCAATATATACC--GAAATTAAGTTCGACGATATAAATAAAGATTTTATTTA 3203  
Qy 3901 GTCTCCAGAAAAAGGGGGAATGAAGACCCACCTGTAGTGTGGCAAGCTAGCTTAAG 3960  
Db 3204 GTCTCCAGAAAAAGGGGGAATGAAGACCCACCTGTAGTGTGGCAAGCTAGCTTAAG 3263  
Qy 3961 TAACGCCATTTTCAAGGCATGG--AAAATACATAACTGAGAAATAGAGAATTCAGATCAA 4019  
Db 3264 TAACGCCATTTTCAAGGCATGGAAAAATACATAACTGAGAAATAGAGAATTCAGATCAA 3323  
Qy 4020 GGTTAGGAACAGA--GAGACAGCAGAAATATGGGCCAAACAGGATATCTGTGTAGCAGTT 4078  
Db 3324 GGTGAGGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGTAGCAGTT 3383  
Qy 4079 CCTGCCCC--GCTCAGGGCCAAAGACAGTTGGAACAGGAGAAATATGGGCCAAACAGGATAT 4137  
Db 3384 CCTGCCCCGCTCAGGGCCAAAGACAGTGAATATGGGCCAAACAGGATAT 3443  
Qy 4138 CTGTGTAAGCAGTTCTGCCCCGCTCAGGGCCAAAGACAGATGGTCCCCAGATGGGT 4197  
Db 3444 CTGTGTAAGCAGTTCTGCCCCGCTCAGGGCCAAAGACAGATGGTCCCCAGATGGGT 3503  
Qy 4198 CCGGCCCTCAGCAGTTCTAGAGAACATCAGATGTTTCCAGGGTGCCTCCAGACCTGA 4257  
Db 3504 CCAGCCCTCAGCAGTTCTAGAGAACATCAGATGTTTCCAGGGTGCCTCCAGACCTGA 3563  
Qy 4258 AATGACCTGTGCTTATTTGAACTAACCAATCAGTTGCTTCTGCTTCTGTTTCGGCG 4317  
Db 3564 AATGACCTGTGCTTATTTGAACTAACCAATCAGTTGCTTCTGCTTCTGTTTCGGCG 3623  
Qy 4318 CTTCGTCTCCCGAGCTCAATAAAGAGCCCAACCCCTCAGTCCGGCGGCCAGTCTCTC 4377  
Db 3624 CTTCGTCTCCCGAGCTCAATAAAGAGCCCAACCCCTCAGTCCGGCGGCCAGTCTCTC 3683  
Qy 4378 CGATAGACTGGTTCGCGGGTACCCGTTTCTCAATAAACCTCTTTCGATTTGCATCCG 4437  
Db 3684 CGATTGACTGAGTGGCGGGTACCCGTTGATCAATAAACCTCTTTCGATTTGCAAAA 3743  
Qy 4438 ACTCGTGTCTGCTGTTCTTGGGAGGGTCTCTCTGAGTGATTTGACTACCCGTCAGGG 4497  
Db 3744 AAA 3803

Qy 4498 GGTCTTTTCAGTTTCTCCACCTACACAGGTCTCTACTAAACATTCCTGATGTGCCGAGGGA 4557  
Db 3804 AAA 3863  
Qy 4558 CTCGTCAGCCCGGTTTTTGTATTATAATAAATCAAGAACAGTGTTCCTTCTCAAGCCAG 4617  
Db 3864 AAA 3923  
Qy 4618 ACTACATCTCTGACTCTCGGCTTTTATAAAGAATTTTGAAGGGCTCTGTGGACTATCTGCC 4677  
Db 3924 AAA 3983  
Qy 4678 ACAGACTTTTAAAGATTTTATGCTCTCTGGATGAGGATTTAGTCAATCTATCTCCTG 4737  
Db 3984 AAA 4043  
Qy 4738 CTATTTTGTGCTGCTCTCCGTTTAAATTTCTAGTTTGCACCTCTTCTCTGAGAGCAC 4797  
Db 4044 AAA 4103  
Qy 4798 GGCATTGCGAGTAGTTAATACTCTGAGGCGAGGCTTCTGTGAAAAGGTTTGCCTGGCT 4857  
Db 4104 AAA 4163  
Qy 4858 CAGTGTGAGATTTTGGCCATAAAGAGGGTCTGCTCCCTGTGTACAGACAGATCGGAATCT 4917  
Db 4164 AAA 4223  
Qy 4918 AGAGTGATATCTCAGAGTCCCGGGTTCGGGCTCTGATCTCAGGCGATCTTTGCTTA 4977  
Db 4224 AAA 4283  
Qy 4978 GAGATCTCTACGCCGCGACGATCGTGGCGGGTACCGAGCTCGAATTCGTAATCATGCT 5037  
Db 4284 AAA 4343  
Qy 5038 CATAGCTGTTTCTGTGTAATTTGTTATCGCTCAATTTCCACACATATACGAGCG 5097  
Db 4344 AAA 4403  
Qy 5098 GAGCATATAAGTGA-----AGCCT 5118  
Db 4404 AAA 4463  
Qy 5119 GGGTGCCTAATGAGTCACTCACTCAATTAATTTGGTGTGCGCTCACTGCCCC----- 5171  
Db 4464 AGNAGCTTAAATGGAGAAAAAATCACTGGATATACCGTTGATATATCCCAATGGC 4523  
Qy 5172 -----GCTTTCCAGTCGGGAAACCTGCTGT 5196  
Db 4524 ATCGTAAAGAACATTTTGGAGCATTTTCAGTCAGTTGCTCAATGTACTATTAACAGACCG 4583  
Qy 5197 GCCAGCTGCAATTAATGAATCGGCCAACCGCGGGGAGAGGGGTTTCGGTATTTCGGGCT 5256  
Db 4584 TTCACTGCTCAATTAATGAATCGGCCAACCGCGGGGAGAGGGGTTTCGGTATTTCGGGCT 4643  
Qy 5257 CTTCGGCTTCTCCTGCTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5316  
Db 4644 CTTCGGCTTCTCCTGCTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4703  
Qy 5317 CAGTCTACTCAAGGGCGGTAAATCGGTTATCCACAGAAATCAGGGGATAAACCGAGAAAGA 5376  
Db 4704 CAGTCTACTCAAGGGCGGTAAATCGGTTATCCACAGAAATCAGGGGATAAACCGAGAAAGA 4763  
Qy 5377 ACATGTAGCAAAAGGCCAGCAAAAGCCAGGAACCGTAAAGCCCGCTTCTGCTGGCT 5436  
Db 4764 ACATGTAGCAAAAGGCCAGCAAAAGCCAGGAACCGTAAAGCCCGCTTCTGCTGGCT 4823  
Qy 5437 TTTTTCATAGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGT 5496  
Db 4824 TTTTTCATAGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGT 4883

5497 GCGGAAACCCGACGACTATAAAGATACACGAGCGTTTCCCGCTGGAAGCTCCCTCGTGC 5556  
5498 GCGGAAACCCGACGACTATAAAGATACACGAGCGTTTCCCGCTGGAAGCTCCCTCGTGC 5557  
5499 GCTCTCTGTTCCGACCCCTGCGCTTACCGGATAACCTGTCGCGCTTCTCCCTTCCGGA 5616  
5500 GCTCTCTGTTCCGACCCCTGCGCTTACCGGATAACCTGTCGCGCTTCTCCCTTCCGGA 5003  
5501 GGTGCGCGCTTCTCATAGCTCAGCTGAGTATCTCAGTTCGGTGTAGTTCGCT 5676  
5502 GGTGCGCGCTTCTCAATGCTCAGCTGAGTATCTCAGTTCGGTGTAGTTCGCT 5063  
5503 CCAAGCTGCGCTGTGTGACGAAACCCCGCTTACGCGCCGACCGCTTATCCGCTA 5736  
5504 CCAAGCTGCGCTGTGTGACGAAACCCCGCTTACGCGCCGACCGCTTATCCGCTA 5123  
5505 ACTATCGTCTGAGTCCAAACCCGCTAAGACACGACTTATCGCCACTGGCAGCAGCCTG 5183  
5506 GTAAACAGGATTACGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGC 5856  
5507 GTAAACAGGATTACGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGC 5184  
5508 CTAACCTACGCTACACTAGAAAGACAGTATTTGGTATCTGCGCTCTGCTGAAAGCCAGTTA 5916  
5509 CTAACCTACGCTACACTAGAAAGACAGTATTTGGTATCTGCGCTCTGCTGAAAGCCAGTTA 5303  
5510 CTTTCGGAAGAAAGTGGTAGCTCTTGATCCGGCAAAACCAACCCGCTGGTAGCGGTG 5363  
5511 CTTTCGGAAGAAAGTGGTAGCTCTTGATCCGGCAAAACCAACCCGCTGGTAGCGGTG 5976  
5512 CTTTCGGAAGAAAGTGGTAGCTCTTGATCCGGCAAAACCAACCCGCTGGTAGCGGTG 5304  
5513 GTTTTTTTTTCGAGGACGAGTATCGGCGAGAAAGGATCTCAAGAGATCCCT 6036  
5514 GTTTTTTTTTCGAGGACGAGTATCGGCGAGAAAGGATCTCAAGAGATCCCT 5423  
5515 TGATCTTTTTCAGCGGCTGACGCTCAGTGGAAACGAAATCTCACGTTAAGGATTTGG 6096  
5516 TGATCTTTTTCAGCGGCTGACGCTCAGTGGAAACGAAATCTCACGTTAAGGATTTGG 5483  
5517 TCATGAGATTATCAAAAAGATCTTCACTAGATCTTTTAAATTAATAAATGAAGTTTA 6156  
5518 TCATGAGATTATCAAAAAGATCTTCACTAGATCTTTTAAATTAATAAATGAAGTTTA 5543  
5519 AATCAATCTAAAGTATATATGATGATAAATTTGGTCTGACAGTTACCAATGCTTAAATCAGTG 6216  
5520 AATCAATCTAAAGTATATATGATGATAAATTTGGTCTGACAGTTACCAATGCTTAAATCAGTG 5603  
5521 AGGCACCTATCTCAGCGATCTGTCTATTTTCATCCATAGTTGCTGACTCCCGCTCG 6276  
5522 AGGCACCTATCTCAGCGATCTGTCTATTTTCATCCATAGTTGCTGACTCCCGCTCG 5604  
5523 TGTAGATAACTACGATAACGAGGCGTTACCATCTGCGCCCGAGTGTCAATGATACCGC 6336  
5524 TGTAGATAACTACGATAACGAGGCGTTACCATCTGCGCCCGAGTGTCAATGATACCGC 5723  
5525 GAGACCCACGCTCAGCGCTCAGATTTATCAGCAATAAACAGCAGCGCGAGGCGCG 6396  
5526 GAGACCCACGCTCAGCGCTCAGATTTATCAGCAATAAACAGCAGCGCGAGGCGCG 5724  
5527 AGCGAGAGTGTCTGCAACTTTATCCGCTCCATCCAGTCTATTATTTGTCGCGG 6456  
5528 AGCGAGCGTGTCTGCAACTTTATCCGCTCCATCCAGTCTATTATTTGTCGCGG 5784  
5529 AAGCTAGAGTAAGTAGTTCGCCAGTTAAAGTTTGGCAACGTTGTCATTTGCTACAG 5903  
5530 AAGCTAGAGTAAGTAGTTCGCCAGTTAAAGTTTGGCAACGTTGTCATTTGCTACAG 6516  
5531 GC-TGCGTGTGTCAGCGCTCGTGTGGTATGGCTTCATTCAGCTCCGCTTCCCAACGAT 6575  
5532 GCATCGTGTGTGTCAGCGCTCGTGTGGTATGGCTTCATTCAGCTCCGCTTCCCAACGAT 5904  
5533 CAAGCGGAGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTAGTCTCCTTCGCTCCTC 6635

5964 CAAGCGAGTTACATGATCCCATGTTGTGCAAAAAGCGGTAGTCTCCTTCGCTCCTC 6023  
6636 CGATGTTGTGACAGAGTAAGTTGGCCGAGTGTATCACTCATGCTTATGGCAGCAGTGC 6695  
6024 CGATGTTGTGACAGAGTAAGTTGGCCGAGTGTATCACTCATGCTTATGGCAGCAGTGC 6083  
6696 ATAAATCTCTTACTGCTCATGTCATCCGATGAGTGTCTTTCTGCTGACTGCTGAGTACTCAA 6755  
6084 ATAAATCTCTTACTGCTCATGTCATCCGATGAGTGTCTTTCTGCTGACTGCTGAGTACTCAA 6143  
6756 CCAAGTCAATCTGAGAAATAGTGTATGCGCGACCGAGTGTCTTTCCCGCGCTCAATAC 6815  
6144 CCAAGTCAATCTGAGAAATAGTGTATGCGCGACCGAGTGTCTTTCCCGCGCTCAATAC 6203  
6816 GGGATAATACCGCGCACATAGCAGAACTTTAAAAGTGTCTCATCATTTGGAACACGTTCTT 6875  
6204 GGGATAATACCGCGCACATAGCAGAACTTTAAAAGTGTCTCATCATTTGGAACACGTTCTT 6263  
6876 GCGGCGGAAAACCTCTCAAGGATCTTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTC 6935  
6264 GCGGCGGAAAACCTCTCAAGGATCTTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTC 6323  
6936 GTGCAACCCAACTGATCTTTGAGCATCTTTTACTTTTCCAGCGCTTTCTGCGTGAGCAAAA 6995  
6324 GTGCAACCCAACTGATCTTTGAGCATCTTTTACTTTTCCAGCGCTTTCTGCGTGAGCAAAA 6383  
6996 GAGGAAGGCAAAATCCGCAAAAAGGAAATAGGCGCACGGAATGTTGAATACTCA 7055  
6384 GAGGAAGGCAAAATCCGCAAAAAGGAAATAGGCGCACGGAATGTTGAATACTCA 6443  
7056 TACTCTCTCTTTTCAATATTTAAGCATTTATCAGGTTATTTGTTCTCATGAGCGGAT 7115  
6444 TACTCTCTCTTTTCAATATTTAAGCATTTATCAGGTTATTTGTTCTCATGAGCGGAT 6503  
7116 ACATATTTGAATGTTTATAGAAAATAAACAATAGGCGTTCGCGCACATTTTCCCGGAA 7175  
6504 ACATATTTGAATGTTTATAGAAAATAAACAATAGGCGTTCGCGCACATTTTCCCGGAA 6563  
7176 AAGTCCGACCTGACGCTCTAAGAACCATTTATCATGACATTAACCTATAAATAAGC 7235  
6564 AAGTCCGACCTGACGCTCTAAGAACCATTTATCATGACATTAACCTATAAATAAGC 6623  
7236 GTATCAGGAGGCGCTTTGCTC 7256  
6624 GTATCAGGAGGCGCTTTGCTC 6644

RESULT 4  
AX823827  
LOCUS AX823827  
DEFINITION Sequence 5 from Patent WO03070958.  
ACCESSION AX823827  
VERSION AX823827.1 GI:39750146  
KEYWORDS murine stem cell retroviral vector (MSCV)  
SOURCE murine stem cell retroviral vector  
ORGANISM murine stem cell retroviral vector  
REFERENCE 1  
AUTHORS Yao,S. and Ellis,J.  
TITLE Retroviral gene therapy vectors including insulator elements to provide high levels of gene expression  
JOURNAL Patent: WO 03070958-A 5 28-AUG-2003;  
THE HOSPITAL FOR SICK CHILDREN (CA)  
FEATURES  
source  
1. .6505  
/organism="murine stem cell retroviral vector"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:258023"  
1. .6505  
/note="n = g, a, t, or c."  
LTR  
ORIGIN

Query Match		42.4%;	Score 3244.4;	DB 6;	Length 6505;
Best Local Similarity		71.1%;	Pred. No. 0;		
Matches 5064;		Conservative 0;	Mismatches 1020;	Indels 1035;	Gaps 25;
Qy	552	TGAAGACCCACCTGTAGTTGGCAAGCTAGCTTAAAGTAAAGCAAGCAATTTTGCAGGCAAT	611		
Db	1	TGAAGACCCACCTGTAGTTGGCAAGCTAGCTTAAAGTAAAGCAATTTTGCAGGCAAT	60		
Qy	612	GGAAATAATCAATCTGAGAAATAGAGAGTTCAGATCAAGGTTAGGAAACAGAGAGACACA	671		
Db	61	GGAAATAATCAATCTGAGAAATAGAGAGTTCAGATCAAGGTTAGGAAACAGAGAGACACA	119		
Qy	672	GAATATGGCCCAACAGAGATATCTGTGTAGCAGTTCCTGCCCGCTCAGGGCCAGAA	731		
Db	120	-----	119		
Qy	732	CAGTTGGAAACAGGAGATATGGCCAAACAGAGATATCTGTGTAGCAGTTCCTGCCCGC	791		
Db	120	-----AGAAATATGGCCAAACAGAGATATCTGTGTAGCAGTTCCTGCCCGC	166		
Qy	792	GCTCAGGCCCAAGAACAGATGGTCCCAAGATGCGGTCCCGCCTCAGCAGTTTCTAGAGA	851		
Db	167	GCTCAGGCCCAAGAACAGATGGTCCCAAGATGCGGTCCCGCCTCAGCAGTTTCTAGAGA	226		
Qy	852	ACCATCAGATGTTTCCAGGGTCCCAAGAGACCTG-AAATGACCTGTGCTTATTTGAA	910		
Db	227	ACCATCAGATGTTTCCAGGGTCCCAAGAGACCTGAAATGACCTGTGCTTATTTGAA	286		
Qy	911	CTAACCAATCAGTTTCGCTTCCTGTTTCGGCGCTTCGTCCCGAGCTCAATAA	970		
Db	287	CTAACCAATCAGTTTCGCTTCCTGTTTCGGCGCTTCGTCCCGAGCTCAATAA	346		
Qy	971	AAGAGCCCAACCCCTCACTCGSGCGCCAGTCTCCGATAGACTGCGTCCCGCGGTA	1030		
Db	347	AAGAGCCCAACCCCTCACTCGSGCGCCAGTCTCCGATAGACTGCGTCCCGCGGTA	406		
Qy	1031	CCCGTATTCCTCAATTAAGCCCTTCGTGTTTGCATCCGAATCGTGAGTCTGCTGATCTT	1090		
Db	407	CCCGTATTCCTCAATTAAGCCCTTCGTGTTTGCATCCGAATCGTGAGTCTGCTGATCTT	466		
Qy	1091	GGAGGGTCTCTCAGATTGATTGACTGCCACACTCGSGGTCTTCAATTTGGAGGTTCC	1150		
Db	467	GGAGGGTCTCTCAGATTGATTGACTGCCACACTCGSGGTCTTCAATTTGGAGGTTCC	526		
Qy	1151	ACCGAGATTTGGAGACCCCTCCAGGAGCACCGACCCCGCCCGGGAGGTAAAGTGG	1210		
Db	527	ACCGAGATTTGGAGACCCCTCCAGGAGCACCGACCCCGCCCGGGAGGTAAAGTGG	586		
Qy	1211	CCAGCACTTATCTGTGTCGTGTC-----CGATTGTCAGTGTCTATGACTGATTTTA	1262		
Db	587	CCAGCGGTCTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTC	646		
Qy	1263	TGCGCTCGCTCGCTAGTACTAGTCTGTATCTGTATCTGTATCTGTATCTGTATCTGTAT	1322		
Db	647	TGCGCTCGCTCGCTAGTACTAGTCTGTATCTGTATCTGTATCTGTATCTGTATCTGTAT	706		
Qy	1323	CTGACGAGTTCGGAACACCCCGCCGCAACCTCGGAGACGTCCTCCAGGACCTTCGGGGCC	1382		
Db	707	CTGACGAGTTCGGAACACCCCGCCGCAACCTCGGAGACGTCCTCCAGGACCTTCGGGGCC	766		
Qy	1383	GTTTTGTGGCCCGACCTGAGTCTCAAAATCCCGATCGTTTGGAGTCTTTTGGTGCACCC	1442		
Db	767	GTTTTGTGGCCCGACCTGAGGAGGGAGTCTGATGTGGAATCCGACCC-----	814		
Qy	1443	CCCTTAGAGGAGGATATGGTTCTGTGTAGGAGACGAGAACCTTAAACAGTTCCTCCGCT	1502		
Db	815	-----CGTCAGGATATGGTTCTGTGTAGGAGACGAGAACCTTAAACAGTTCCTCCGCT	867		
Qy	1503	CCGTCTGAATTTTGTCTTGGTTCGGAACCGAGCCGCGCGCTTGTGTTTGTGCTGCTGCGGC	1559		
Db	868	CCGTCTGAATTTTGTCTTGGTTCGGAACCGAGCCGCGCGCTTGTGTTTGTGCTGCTGCGGC	927		
Qy	1560	TGCAGCATCGTTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGT	1619		

Db	928	TGCAGCATCGTTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGT	987		
Qy	1620	GGCCCGGCTAGACTGTTTACCACCTCCCTTAAAGTTTGAAGTTCAGTTCAGTTCAGTTCAGT	1679		
Db	988	GGCC-----AGACTGTTTACCACCTCCCTTAAAGTTTGAAGTTCAGTTCAGTTCAGTTCAGT	1041		
Qy	1680	GAGCGGATCGCTCAACACAGTTCGATGTCAAGAAAGAGACGTTGGTTTACCTTCGCTTCG	1739		
Db	1042	GAGCGGATCGCTCAACACAGTTCGATGTCAAGAAAGAGACGTTGGTTTACCTTCGCTTCG	1101		
Qy	1740	TCTCAGAAATGGCCCAACCTTTAAAGTTCGATGGCGCGGAGACGCGACCTTTAAACCGAGAC	1799		
Db	1102	TCTCAGAAATGGCCCAACCTTTAAAGTTCGATGGCGCGGAGACGCGACCTTTAAACCGAGAC	1161		
Qy	1800	CTCATATCCACAGGTTTAAAGTCAAGGTCTTTTCACTGGCCCGCATGGACACCCAGACCCAG	1859		
Db	1162	CTCATATCCACAGGTTTAAAGTCAAGGTCTTTTCACTGGCCCGCATGGACACCCAGACCCAG	1221		
Qy	1860	GTCCCTTACATCGTGAAGCTGGGAGCTTGGCTTTTGAACCCCTCCCTGGGTCAAGGCC	1919		
Db	1222	GTCCCTTACATCGTGAAGCTGGGAGCTTGGCTTTTGAACCCCTCCCTGGGTCAAGGCC	1281		
Qy	1920	TTTGTACACCTAAGCCTTCGCTCTCTTCTCCTCATCGCGCCGCTCTCTCCCTTCGAA	1979		
Db	1282	TTTGTACACCTAAGCCTTCGCTCTCTTCTCCTCATCGCGCCGCTCTCTCCCTTCGAA	1341		
Qy	1980	CCTCTCTGTTTCGACCCCGCTCGATCTCTCCCTTTATCCAGCCCTCAGTCTCTCTAGGC	2039		
Db	1342	CCTCTCTGTTTCGACCCCGCTCGATCTCTCCCTTTATCCAGCCCTCAGTCTCTCTAGGC	1401		
Qy	2040	GCCCCATATGGCCATATGAGATCTTATATGGGGCACCCCGCCCTTTGTAACCTTCCT	2099		
Db	1402	GCCGGAAT-----	1423		
Qy	2100	GACCTGACATGACAAAGATTACTTAAACGCCCTCTCTCAAGCTCTACTTACAGGCTTCT	2159		
Db	1424	AGATCTAATTTACTACCGGTAGGGAGGCGCTTTTCCCAAGGACGCTCTGGAGCATGCGCTT	1483		
Qy	2160	ACTTAGTTCAGACGAAATCTGGAGACTCTGCGCGGACGCTACCAAGAAACAACTGGACC	2219		
Db	1484	TAGCAGCCCGCTGGCATTGGCGCTACAAAGTGGCTCTGGCTCGCACATTCAC	1543		
Qy	2220	GACCGGTGTTACCTTACCGAGTTCGCGGACACAGTGTGGGTTCGCGACACACAGA	2279		
Db	1544	ATCCACCGGTAGCGC-----AACCGGCTCGCTTCTTGTGGCCCTCTCGCGCA	1594		
Qy	2280	CTAAGAACCTTAGAACCTCGCTGGTGGAAAGACCTTACAGTCTCTGTTGACCAACCCCA	2339		
Db	1595	CTTCTACTCC-----TCCCTTAGTCAAGAAATTTCCCGCAGCAAGCTCGCGTCTGCAGGA	1650		
Qy	2340	CCCTCAAGTAGACGGCATCGCAGCTTGGATACACGCGCCCGCCACGTCGAGGCTGCCGACC	2399		
Db	1651	CGTGACAAATGGAAGTAGCAGCTCTACTAGTCTCTGTCAGATGGAACGACCGCTGAGC	1710		
Qy	2400	CCGGGGTGGACCACTCTCTAGACTGCGCATGGGATGGAGCTGTATCTCTCTCTCTCTCT	2459		
Db	1711	AATGGAAGCGGTAGGCTTTGGGGCAGCGCCAAATAGCAGCTTGTCTCTCTCTCTCTCT	1770		
Qy	2460	AGCAACAGCTTACAGTGTCTCACTCCGACATCCAGTGTACCCAGACCCAGACGCTGAG	2519		
Db	1771	GGGTCTAGAGGCTGGGAAGGGGTGGGTCCGGGGCGGGCTCAG-----	1813		
Qy	2520	CGCCAGCGTGGTGACAGTAGTACCATCACTGTGTAGGCCAGTCAAGATGTTGGGTACTTTC	2579		
Db	1814	-----GGGCGGGCTCAGGGCGGGCGGGCGCCGAAAGTCTCTCCGAGGCGCGGATCT	1868		
Qy	2580	TGTAGCTTGTGTACAGCAGAAAGCAGGTAAAGCTTCAAAAGCTGTGATCTACTTGACATC	2639		
Db	1869	GCAGCTTCAAAAGCGCAGCTGTGCGCGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1928		
Qy	2640	CACCCGCGACACTGTTGTGTCAGACAGATTACGCGGTAGCGGTAGCGGTACCGACTTCAC	2699		

Db 1929 CGACCTGCAGCCAAATATATGGGATCGGCCATTTGAACAAGATGGATGCAACGAGTTCTC-- 1986  
Qy 2700 CTTTACCATCAGAGCTCCAGCAGAGACATCGCCACTACTACTGCGCAGCAATATAG 2759  
Db 1987 -----CGGCCGCTTGGGTGGAGAGGCTAT--TCGGCTATGACTGGGCAACAAGAC 2035  
Qy 2760 CCTTATCGGTGCTTGGCCAAAGGAGCAAGGTGGAAATCAAAAGAGGTGGCTCAGGATC 2819  
Db 2036 ATTCGGCTGCTC-----TGATGCCCGGTGT 2062  
Qy 2820 GGGTGGATCCGGCTCTGTGGGTCTCAGGATCGGAGTCCAACTGGTGGAGAGCGGTGGAGG 2879  
Db 2063 CCGGCTGTTCAGCGCAGGGGCGCCGGTCTCTTTTGTCAAGACCGACCTGTCCGGTGCCT 2122  
Qy 2880 TGTGTGCAACTGGCCGCTCCCTGGCTGTCTCTCGCATCTGGCTTGGCTTGGATTCAC 2939  
Db 2123 GAATGAACATGCAAGGACGAGCGAGCGGCTATCGTGGCTGGCCACGAGCGGGCTTCTT 2182  
Qy 2940 CACATATTGGATGAGTTGGGTGAGACAGGCACCTCGAAAGGTCTTGGTGGATTTGGAGA 2999  
Db 2183 CGCAGCTGTCTCGACGTTGTCACTGAAGCGGGAAGGACTGGCTGCTATTGGCGAAGT 2242  
Qy 3000 AATTCAATCCAGATAGCAGTAACTAATATGCGGC-GTCTCTAAGAGATAGATTACAA 3058  
Db 2243 GCGGGGAGGATCTCTGTCACTCACCTTGTCTCGCGAGAAAGTATCCATCATGCG 2302  
Qy 3059 TATCGGAGACACCGCAAGAACATTTCTTCTCAATGACAGCCTGACACCGAAG 3118  
Db 2303 TGATGCAATGCGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCAACAGC 2362  
Qy 3119 ACACCGGGTCTATTTTGTGCAAGCTTTACTTGGCTTCCTCGTGTGCTTGTATTGGG 3178  
Db 2363 GAAC-----ATCGATCGAGCAGCAGTACTGGATGGAGCGGTCTTGTGCATGAGA 2419  
Qy 3179 GCCAAGGACCCCGTCACTGCTCCAGTCTCAAGCTTAAGCCCAACAGCAGCGCCGCGAC 3238  
Db 2420 TGATCTGGACGAAGACATCAGGGCTCGCGCAGCGGAATGTTTCGCCAGGCTCAAGC 2479  
Qy 3239 CACCAACACCGGGCCCATCGGTGCGAGCCCTGTCCCTGCGCCAGAGGGGCTC 3298  
Db 2480 GCGCATGCCCAGCGGAGGATCTCGTGTGATCCCATGGC----- 2519  
Qy 3299 GGCACGCGCGGGGCGCAGTGCAACAGAGGGGCTGGACTTCGCCCTGGATCCCAAC 3358  
Db 2520 -----GATGCTGCT 2529  
Qy 3359 TGTGTACTCTGTGATGAATCTCTTCACTATGTTGTCATTTCTACTGCTGTTC 3418  
Db 2530 TGCCGAATATCATGTGGAATAATGCGGCTTTTCTGGATTCATCGACTGTG----- 2580  
Qy 3419 TGAGAGTGAAGTTACGAGGAGCGGAGAGCCCCCGGTACCGAGGCGCCAGACCAGC 3478  
Db 2581 -----GCGGCTGGGTGTGGGAGCC 2601  
Qy 3479 TCTATAACGAGCTCAATCTAGGACGAAGAGAGGATGATGTTTGGACAAGAGACGTG 3538  
Db 2602 GCTATCAGG----- 2610  
Qy 3539 GCGGAGCCCTGAGATGGGGGGAAGCCGAGAAGGAACCCCTCAGGAAGGCTGTACA 3598  
Db 2611 ----- 2610  
Qy 3599 ATGAATGCAGAAAGATAGATGGCGGAGCCCTACAGTGAATGGGATGAAGGCGAGC 3658  
Db 2611 -----ACATAGCGTTGGCTACCCGCTGATATGCTGAAGAGCTTGGCG 2652  
Qy 3659 GCGGAGGGGCAAGGGCAGATGGCTTTTACAGGGTCTCAGTACAGCCACCAAGGACA 3718  
Db 2653 GCGAATGGGTGACCGCTCTCTGTGCTTTACGGTATCGCGCTCCGATTCGACGCGCA 2712  
Qy 3719 CTTACGACCCCTTCACTACGAGGCCCTGCCCCCTCGCTAACTCGACGCGCGCGGATC 3778  
Db 2713 TCGCCTTCTATCGCCTTCTTGACG----- 2736

Qy 3779 CGGATTAGTCCAAATTTGTTAAAGACAGGATATCAGTGTCCAGGCTCTAGTTTTCATCA 3838  
Db 2737 -----AGTTCTTCTGAGGGATCGTTCGACCTGCGAGCAAGCTT----- 2775  
Qy 3839 ACAATATCACGAGCTGAAGCCTATAGATGACGAGCATAGATAAATAAAGATTTTATT 3898  
Db 2776 -----ATCGATAAAATAAAGATTTTATT 2799  
Qy 3899 TAGTCTCCAGAAAGGGGGAATGAAGACCCCACTGTAGTTTGGCAGCTAGCTTAA 3958  
Db 2800 TAGTCTCCAGAAAGGGGGAATGAAGACCCCACTGTAGTTTGGCAGCTAGCTTAA 2859  
Qy 3959 AGTAAACGCATTTTCAAGGCGATGAAAATACATAACTGAGAAATAGAGAAAGTTTCAGATCA 4018  
Db 2860 AGTAACGCCATTTTCAAGGCGATGAAAATACATAACTGAGAAATAGAGAAAGTTTCAGATCA 2919  
Qy 4019 AGTTAGAAACAGAGAGACAGCAGAAATATGGGCCAAACAGGATATCTGTGTGAAGAGTT 4078  
Db 2920 AGTTTAGGAACAGAGAGACAGC----- 2941  
Qy 4079 CTTGCCCGCTCAGGCGCCAGAACAGTTTGGAAACAGGAGAAATATGGCCAAACAGGATATC 4138  
Db 2942 -----AGAATATGGCCAAACAGGATATC 2965  
Qy 4139 TGTGTAAGCAGTTCTCTGCCCGCTCAGGGCCAAAGAACAGATGTTCCCGATGCGGTC 4198  
Db 2966 TGTGTAAGCAGTTCTCTGCCCGCTCAGGGCCAAAGAACAGATGTTCCCGATGCGGTC 3025  
Qy 4199 CCGCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTCCCGCAAGAGCTG-A 4257  
Db 3026 CCGCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTCCCGCAAGAGCTGAA 3085  
Qy 4258 AATGACCTGTGCTTATTGAACTAAACCAATCAGTTGCTTCTCGCTTCTGTTCCGCG 4317  
Db 3086 AATGACCTGTGCTTATTGAACTAAACCAATCAGTTGCTTCTCGCTTCTGTTCCGCG 3145  
Qy 4318 CTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGCGCGCAGTCTCTC 4377  
Db 3146 CTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGCGCGCAGTCTCTC 3205  
Qy 4378 CGATAGACTGCTGCTCGCCCGGTACCCGTGTTCTCAATAAACCCCTCTTGAGTTGATCCG 4437  
Db 3206 CGATAGACTGCTGCTCGCCCGGTACCCGTGTTCTCAATAAACCCCTCTTGAGTTGATCCG 3265  
Qy 4438 ACTCGTGTCTGCTGTTCTCTGAGGGGTCT-CTCTGAGTGAATGACTACCGGTCAGCG 4496  
Db 3266 ACTCGTGTCTGCTGTTCTCTGAGGGGTCTCTCTGAGTGAATGACTACCGGTCAGCG 3325  
Qy 4497 GGGTCTTTTCAAGTTTCTCCACCTACACAGGTCTCACTAAACATTCCTGATGTGCCG 4556  
Db 3326 GGGTCTT----- 3333  
Qy 4557 ACTCCGTCAGCCCGGTTTTTGTGTTATAATAAATAAAGAACAGTGTCCCTTCAAGCCA 4616  
Db 3334 ----- 3333  
Qy 4617 GACTACATCCTGACTCTCGGCTTTATAAAGAAATGTTGAGGGGCTCTGTGGATATCTGC 4676  
Db 3334 ----- 3333  
Qy 4677 CACACGACTTTTAAAGATTTTATGCTCTCGATGAGGATTTAGTCAATCTATCTCTCG 4736  
Db 3334 ----- 3333  
Qy 4737 TCTATTTTGTGGCTCTCGTATTTTAAATTTTCTAGTTTGCATCTCCCTTCTGTGAGACA 4796  
Db 3334 ----- 3333  
Qy 4797 CCGGATTCAGAGTAGTTAATCTCTGAGGCGAGGCTTCTGTGAAAGGTTTCCCTGGGC 4856  
Db 3334 ----- 3333

QY 4857 TCAGTGTGAGATTTTGCCATAAAAGGGGTCTGCCCCCTGTGTACAGACATCGGAATC 4916  
Db 3334 -----TCAANN 3361  
QY 4917 TAGAGTGCACTACAGAGTCCCGCGGTTCGGGGCTCTGATCTCAGGGCATCTTTGGCT 4976  
Db 3362 NNN 3421  
QY 4977 AGAGATCCTTACGCCGAGCGATCGTGGCGGGTACCGAGCTCGAATTCGTAATCATGG 5036  
Db 3422 NNN 3481  
QY 5037 TCATAGCTGTTCTCTGTGTGAAATGTTATCGCTCACAATTCACACAAATACGAGCC 5096  
Db 3482 TCATAGCTGTTCTCTGTGTGAAATGTTATCGCTCACAATTCACACAAATACGAGCC 3541  
QY 5097 GGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAAATGAGTGAGTAACTCACAATTAATTCG 5156  
Db 3542 GGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAAATGAGTGAGTAACTCACAATTAATTCG 3601  
QY 5157 TTGGGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTCGTCAGCTGCATTAATGAATC 5216  
Db 3602 TTGGGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTCGTCAGCTGCATTAATGAATC 3661  
QY 5217 GGCCAAACGCGGGAGAGCGGTTTGGTATTTGGCGGCTCTTCGCTTCCTCGCTCACT 5276  
Db 3662 GGCCAAACGCGGGAGAGCGGTTTGGTATTTGGCGGCTCTTCGCTTCCTCGCTCACT 3721  
QY 5277 GACTCGCTGCGCTCGTTCGCTGCGGCGAGCGGTATCAGCTCACTCAAGAGCGGTA 5336  
Db 3722 GACTCGCTGCGCTCGTTCGCTGCGGCGAGCGGTATCAGCTCACTCAAGAGCGGTA 3781  
QY 5337 ATACGGTTATCCACAGATCAGGGGATTAACGAGGAAAGACATGTAGCAAAAGGCCAG 5396  
Db 3782 ATACGGTTATCCACAGATCAGGGGATTAACGAGGAAAGACATGTAGCAAAAGGCCAG 3841  
QY 5397 CAAAAGGCCAGAACCGTAAAGAGCGGTTGCTGGGTTTTTCCATAGGCTCGGCC 5456  
Db 3842 CAAAAGGCCAGAACCGTAAAGAGCGGTTGCTGGGTTTTTCCATAGGCTCGGCC 3901  
QY 5457 CCTCAGCAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGGAAACCCGACAGGACTA 5516  
Db 3902 CCTCAGCAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGGAAACCCGACAGGACTA 3961  
QY 5517 TAAAGATACCAAGGGTTTTCCCTCGGAGCTCCCTCGTGCCTCTCTGTTCCGACCCCTG 5576  
Db 3962 TAAAGATACCAAGGGTTTTCCCTCGGAGCTCCCTCGTGCCTCTCTGTTCCGACCCCTG 4021  
QY 5577 CCGCTTACCGGATACCTGTCCGCTTTCTCCCTCGGGAAGCGTGGGCTTTCTCATAGC 5636  
Db 4022 CCGCTTACCGGATACCTGTCCGCTTTCTCCCTCGGGAAGCGTGGGCTTTCTCATAGC 4081  
QY 5637 TCAGCTGTAGTATCTCAGTTCGGTGTAGTCTCGTCCGCTCAAGCTGGGCTGTGTCAC 5696  
Db 4082 TCAGCTGTAGTATCTCAGTTCGGTGTAGTCTCGTCCGCTCAAGCTGGGCTGTGTCAC 4141  
QY 5697 GAAACCCCGCTTACGCCGACCGCTGCGCTTTATCCGGTAACTATCGTCTTGAAGTCCAAC 5756  
Db 4142 GAAACCCCGCTTACGCCGACCGCTGCGCTTTATCCGGTAACTATCGTCTTGAAGTCCAAC 4201  
QY 5757 CCGGTAAAGACAGCATTAATCCCACTGGCAGCAGCACTGTGTAAACAGGATAGCAGAGCG 5816  
Db 4202 CCGGTAAAGACAGCATTAATCCCACTGGCAGCAGCACTGTGTAAACAGGATAGCAGAGCG 4261  
QY 5817 AGGTATGTAGCGGTGTACAGATTTCTTGAAGTGGTGGCTTAACCTACCGGTACACTAGA 5876  
Db 4262 AGGTATGTAGCGGTGTACAGATTTCTTGAAGTGGTGGCTTAACCTACCGGTACACTAGA 4331  
QY 5877 AGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAGAGTTGGT 5936  
Db 4322 AGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAGAGTTGGT 4381  
QY 5937 AGCTCTTGATCCGGCAAAACCAACCCGCTGCTAGCGGTGTTTTTTTGTGCAAGCAG 5996

Db 4382 AGCTCTTGATCCGGCAAAACAAACCCGCTGCTAGCGGTGTTTTTTGTTGCAAGCAG 4441  
QY 5997 CAGATTACGCGCAGAAAAAGAGTCTCAAGAAATCCTTTGTATCTTTCTACGGGGTCT 6056  
Db 4442 CAGATTACGCGCAGAAAAAGAGTCTCAAGAAATCCTTTGTATCTTTCTACGGGGTCT 4501  
QY 6057 GAGCCTCAGTGGAAACGAAAACTCAGTTAAGGGATTTTGGTTCATGAGATTAATCAAAAAAG 6116  
Db 4502 GAGCCTCAGTGGAAACGAAAACTCAGTTAAGGGATTTTGGTTCATGAGATTAATCAAAAAAG 4561  
QY 6117 ATCTTCACTAGATCCTTTTAAATTAATAATGAAGTTTAAATCAATCAATGAATATAT 6176  
Db 4562 ATCTTCACTAGATCCTTTTAAATTAATAATGAAGTTTAAATCAATCAATGAATATAT 4621  
QY 6177 GAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACTTATCTCAGCGATC 6236  
Db 4622 GAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACTTATCTCAGCGATC 4681  
QY 6237 TGTCTATTTGTTTCAATCATAAGTTGCTGCTCCCGTCTGTTAGATAAATCAGATACGG 6296  
Db 4682 TGTCTATTTGTTTCAATCATAAGTTGCTGCTCCCGTCTGTTAGATAAATCAGATACGG 4741  
QY 6297 GAGGGCTTACCATCTGGCCCCAGTCTGCAATGATACCGCAGACCCAGCTCACCGCT 6356  
Db 4742 GAGGGCTTACCATCTGGCCCCAGTCTGCAATGATACCGCAGACCCAGCTCACCGCT 4801  
QY 6357 CCAGATTATCAGCAATAAACCAGCCAGGAGGCGGAGCGCAGAAAGTGGTCTCTCA 6416  
Db 4802 CCAGATTATCAGCAATAAACCAGCCAGGAGGCGGAGCGCAGAAAGTGGTCTCTCA 4861  
QY 6417 ACTTTATCCGCTCCATCCAGTCTTAATTTGTTGCGGGAGCTAGAGTAGTTCG 6476  
Db 4862 ACTTTATCCGCTCCATCCAGTCTTAATTTGTTGCGGGAGCTAGAGTAGTTCG 4921  
QY 6477 CCAGTTAATAGTTTGGCAAGCTTTGTCCTATGCTACAGCC - TCGTGGTGTACGCTCG 6535  
Db 4922 CCAGTTAATAGTTTGGCAAGCTTTGTCCTATGCTACAGCCATCGTGGTGTACGCTCG 4981  
QY 6536 TCGTTTGGTATGCTTCAATTCAGCTCCGCTTCCCAACGATCAAGCGAGTTACATGATCC 6595  
Db 4982 TCGTTTGGTATGCTTCAATTCAGCTCCGCTTCCCAACGATCAAGCGAGTTACATGATCC 5041  
QY 6596 CCAATGTTGTGCAAAAAAGCGTTAGCTCTTCGCTCCCGATCGTTGTGTCAGAGTAAG 6655  
Db 5042 CCAATGTTGTGCAAAAAAGCGTTAGCTCTTCGCTCCCGATCGTTGTGTCAGAGTAAG 5101  
QY 6656 TTGGCCGCGAGTTTATCCTCATGCTTATGGCAGCAGCTGCAATAATCTCTTACTGTCTAG 6715  
Db 5102 TTGGCCGCGAGTTTATCCTCATGCTTATGGCAGCAGCTGCAATAATCTCTTACTGTCTAG 5161  
QY 6716 CCATCCGTAAGATGCTTTTCTGCTGCTGCTGAGTACTCAACCAAGTCAATCTTGAGAAATAG 6775  
Db 5162 CCATCCGTAAGATGCTTTTCTGCTGCTGCTGAGTACTCAACCAAGTCAATCTTGAGAAATAG 5221  
QY 6776 TGATGCGGGACCGAGTTGCTCTTTCGCGCGCTCAATACGGGATAATACCGCGCCACAT 6835  
Db 5222 TGATGCGGGACCGAGTTGCTCTTTCGCGCGCTCAATACGGGATAATACCGCGCCACAT 5281  
QY 6836 AGCAGAACTTTTAAAGTGTCTCATCATTTGGAAAAAGCTTCTTCGGGGCGAAAACTCTCAAGG 6895  
Db 5282 AGCAGAACTTTTAAAGTGTCTCATCATTTGGAAAAAGCTTCTTCGGGGCGAAAACTCTCAAGG 5341  
QY 6896 ATCTTACCGCTGTGAGATCCAGTTCGATTAACCCACTCGTGCACCAACTGATCTTCA 6955  
Db 5342 ATCTTACCGCTGTGAGATCCAGTTCGATTAACCCACTCGTGCACCAACTGATCTTCA 5401  
QY 6956 GCATCTTTTACTTTTACAGCGTTTCTGGGTGAGCAAAAAACAGGAAGCAAAATGCGCGCA 7015  
Db 5402 GCATCTTTTACTTTTACAGCGTTTCTGGGTGAGCAAAAAACAGGAAGCAAAATGCGCGCA 5461  
QY 7016 AAAAAGGAATAAGGGCGCACCGAAATGTTGAATACTCATACTCTCTCTTTTCAATAT 7075



Db 5462 AAAAAAGGGAATAAGGCGCAGCAGGAAATGTTGAATATCTCATATCTCTCTCTTCAATAT 5521  
 QY 7076 TATTGAAGCATTTATCAGGGTTATTCTCTCATGAGCGGATACATATTTCGAATGATTTTAG 7135  
 Db 5522 TATTGAAGCATTTATCAGGGTTATTCTCTCATGAGCGGATACATATTTCGAATGATTTTAG 5581  
 QY 7136 AAAAAATAAACAATAGGGGTTCCGCGCACATTTCCCGGAAAGTGCACCTGACGCTCTAA 7195  
 Db 5582 AAAAAATAAACAATAGGGGTTCCGCGCACATTTCCCGGAAAGTGCACCTGACGCTCTAA 5641  
 QY 7196 GAAACCATTTATCATGATGACATTAACCTATAAAATAGCGGTATACAGAGCGCTTTCTGT 7255  
 Db 5642 GAAACCATTTATCATGATGACATTAACCTATAAAATAGCGGTATACAGAGCGCTTTCTGT 5701  
 QY 7256 CTCGCGCGTTTCGGGTGATCAGCGTGAACCTCTGACACATCGAGCTCCCGGAGAGGTC 7315  
 Db 5702 CTCGCGCGTTTCGGGTGATCAGCGTGAACCTCTGACACATCGAGCTCCCGGAGAGGTC 5761  
 QY 7316 ACAGCTTGTCTGTAAAGCGATGCGGGAGCAGACAAAGCCGTCAGGGCGCGTCAGCGGT 7375  
 Db 5762 ACAGCTTGTCTGTAAAGCGATGCGGGAGCAGACAAAGCCGTCAGGGCGCGTCAGCGGT 5821  
 QY 7376 GTTGGCGGTGTCGGGGCTGGCTTAACATATGCGGCATCAGAGCAGATTTGATCTGAGAGTG 7435  
 Db 5822 GTTGGCGGTGTCGGGGCTGGCTTAACATATGCGGCATCAGAGCAGATTTGATCTGAGAGTG 5881  
 QY 7436 CACCATATCGGTGTGAATATCGGCACAGATCGTAAAGAGAAATACCGCATCAGGCGC 7495  
 Db 5882 CACCATATCGGTGTGAATATCGGCACAGATCGTAAAGAGAAATACCGCATCAGGCGC 5941  
 QY 7496 CATTCGCCATTACGGCTGCGCAACTGTGGGAAGGCGCATCGGTGGGGCTCTTCGCTA 7555  
 Db 5942 CATTCGCCATTACGGCTGCGCAACTGTGGGAAGGCGCATCGGTGGGGCTCTTCGCTA 6001  
 QY 7556 TTACGCCAGCTGCGGAAAGGGGATGTCTGCAAGGCGATTAAGTTGGGTAAAGCAGG 7615  
 Db 6002 TTACGCCAGCTGCGGAAAGGGGATGTCTGCAAGGCGATTAAGTTGGGTAAAGCAGG 6061  
 QY 7616 TTTTCCCACTCAGAGTGTAAACGAGCGCCAGTGCC 7654  
 Db 6062 TTTTCCCACTCAGAGTGTAAACGAGCGCCAGTGCC 6100

RESULT 5  
 LOCUS CQ879092 5903 bp DNA linear PAT 11-OCT-2004  
 DEFINITION Sequence 5 from Patent WO2004062554.  
 ACCESSION CQ879092  
 VERSION CQ879092.1 GI:54033469  
 KEYWORDS synthetic construct  
 SOURCE synthetic construct  
 ORGANISM other sequences; artificial sequences.  
 REFERENCE 1  
 AUTHORS Martinez-Serrano,A., Liste,I. and Villa,A.  
 TITLE Enhancement of neuron generation and survival  
 JOURNAL Patent: WO 2004062554-A 5 29-JUL-2004;  
 NSGENE A/S (DK)

FEATURES  
 source Location/Qualifiers  
 1..5903  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="Viral vector"

ORIGIN  
 Query Match 42.1%; Score 3223.6; DB 6; Length 5903;  
 Best Local Similarity 73.9%; Pred. No. 0;  
 Matches 4964; Conservative 0; Mismatches 759; Indels 995; Gaps 29;

QY 550 TTTGAAAGACCCACCTGTAGGTTTCGCAAGCTAGCTTAAGTAAACGCCATTTTGCAGGC 609  
 Db 167 TGTGAAAGACCCACCTGTAGGTTTCGCAAGCTAGCTTAAGTAAACGCCATTTTGCAGGC 226

QY 610 ATGG-AAAATACATACTAGAAAGTTTCAATCAAGTTAGGAAACAGA-GAGAC 667  
 Db 227 ATGGAATAATACATACTAGAAAGTTTCAATCAAGTTAGGAAACAGA-TGAAC 286  
 QY 668 AGCAAAATATGGGCAAAACAGAGATATCTGTGTAAAGAGTTCCTGCCCC-GCTCAGGGCC 726  
 Db 287 AGCTGAATATGGGCAAAACAGCGATATCTGTGTAAAGAGTTCCTGCCCCGCTCAGGGCC 346  
 QY 727 AAGAACAGTTTGAACAGAGAGATATGGGCAAAACAGGATATCTGTGTAAAGAGTTCCTG 786  
 Db 347 AAGAACAGATGGAACAGCTGATATGGGCAAAACAGGATATCTGTGTAAAGAGTTCCTG 406  
 QY 787 CCCCCTGCTCAGGCAAAACAGATGTTCCCAAGATGCGGTCCCGCTCCAGCAGTTTCT 846  
 Db 407 CCCCCTGCTCAGGCAAAACAGATGTTCCCAAGATGCGGTCCCGCTCCAGCAGTTTCT 466  
 QY 847 AGAAGAACCATCAGATGTTTCCAGGGTCCCAAGAGCTGAAATGACCTGTGCTTAT 906  
 Db 467 AGAAGAACCATCAGATGTTTCCAGGGTCCCAAGAGCTGAAATGACCTGTGCTTAT 526  
 QY 907 TGAACCTAACCATCAGATGTTTCCAGGGTCCCAAGAGCTGAAATGACCTGTGCTTAT 966  
 Db 527 TGAACCTAACCATCAGATGTTTCCAGGGTCCCAAGAGCTGAAATGACCTGTGCTTAT 586  
 QY 967 ATAAAGAGAGCCCAACACCTCCTCCTCGGCGCCAGTCTCTCGATAGACTGCTCGCCCG 1026  
 Db 587 ATAAAGAGAGCCCAACACCTCCTCCTCGGCGCCAGTCTCTCGATAGACTGCTCGCCCG 646  
 QY 1027 GGTACCCCTGATTTCCCAATAAAGCTCTTCTGTGTTTGCATCCGAATCGTGGACTCGCTGAT 1086  
 Db 647 GGTACCCCTGATTTCCCAATAAAGCTCTTCTGTGTTTGCATCCGAATCGTGGACTCGCTGAT 705  
 QY 1087 CTTTGGAGGGTCTCTCAGATTTGATTTGACTGCCCC-ACCTCGGGGGTCTTTTCAATTGGAG 1145  
 Db 706 CTTTGGAGGGTCTCTCAGATTTGATTTGACTGCCCCGCTCAGCGGGGTCTTTTCAATTGGGG 765  
 QY 1146 GTTCCACCGAGATTTGGAGACCCCTGCGGAGACCCAGGACCCCGCCCGGGAGGTAA 1205  
 Db 766 GCTGCTCGGGATCGGGAGACCCCTGCGGAGACCCAGGACCCAGGACCCCGGGAGGTAA 824  
 QY 1206 GCTGCGCCAGCAACTTATCTGTGTTTCCGATTTGATGTTGATGATGATTTTATGTC 1265  
 Db 825 GCTGCGCCAGCAACTTATCTGTGTTTCCGATTTGATGTTGATGATGATTTTATGTC 884  
 QY 1266 GCTGCGGTGCTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 1325  
 Db 885 GCTGCGGTGCTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 944  
 QY 1326 ACGAGTTTGGAAACACCGGCGCAACCTTGGGAGAGCTCCAGGAGCTTCCGGGGGCGGTT 1385  
 Db 945 ACGAGTTTGGAAACACCGGCGCAACCTTGGGAGAGCTTCCAGGAGCTTCCGGGGGCGGTT 1004  
 QY 1386 TTTTGGCCCCGACCTGAGTCTTAAATCCCGATCGTTTATAGGACTCTTTGGTGACCCCCC 1445  
 Db 1005 TTTTGGCCCCGACCTGAGTCTTAAATCCCGATCGTTTATAGGACTCTTTGGTGACCCCCC 1064  
 QY 1446 TTAGAGGAGGGATATGTTGTTCTGTTAGAGAGAGAGAACTTAAACAGTTTCCCGCTCCG 1505  
 Db 1065 TTAGAGGAGGGATATGTTGTTCTGTTAGAGAGAGAGAACTTAAACAGTTTCCCGCTCCG 1124  
 QY 1506 TCTGAATTTTTCGTTTTCGTTTGGGACCGAAGCGCGCGCTCTGCTGCTGCTGCTGCTGCT 1565  
 Db 1125 TCTGAATTTTTCGTTTTCGTTTGGGACCGAAGCGCGCGCTCTGCTGCTGCTGCTGCTGCT 1184  
 QY 1566 ATCGTTTCTGTTGTTCTGTTGTTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTT 1625  
 Db 1185 ATCGTTTCTGTTGTTGTTGTTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTT 1244  
 QY 1626 GCGTAGACTGTTTACACTCCCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTT 1685  
 Db 1245 GCGTAGACTGTTTACACTCCCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTT 1304





Db 2890 GCG-----GGACTCTGGGTTTCGATAAAATAAAGATTTTTATTTAGTCT 2933

Qy 3905 CCAGAAAAAGGGGGAATGAAGACCCCACTGTAGTTTGGCAAGCTTAACTAAC 3964

Db 2934 CCAGAAAAAGGGGGAATGAAGACCCCACTGTAGTTTGGCAAGCTTAACTAAC 2993

Qy 3965 GCATTTTTCAGGCGATGG-AAAATACATAAATCTGAGAATAGAGAAGTTTCAGATCAAGGTT 4023

Db 2994 GCATTTTTCAGGCGATGG-AAAATACATAAATCTGAGAATAGAGAAGTTTCAGATCAAGGTT 3053

Qy 4024 AGGAACAGA-GAGACAGCAGAAATATGGGCCAAACAGGATATCTGTGTAAGCAGTTTCCGTG 4082

Db 3054 AGGAACAGATGGAAACAGCTGAATATGGGCCAAACAGGATATCTGTGTAAGCAGTTTCCGTG 3113

Qy 4083 CCCC-GCTCAGGCGCAAGAACAGTTTGGAAACAGAGAAATATGGGCCAAACAGGATATCTGTG 4141

Db 3114 CCCCCTCAGGCGCAAGAACAGATGGAAACAGCTGAATATGGGCCAAACAGGATATCTGTG 3173

Qy 4142 GGTAAAGCAGTTCTGCTCCCGGCTCAGGGCCCAAGAACAGATGGTCCCGAGATGGGTCCCG 4201

Db 3174 GGTAAAGCAGTTCTGCTCCCGGCTCAGGGCCCAAGAACAGATGGTCCCGAGATGGGTCCCG 3233

Qy 4202 CCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCAGGACCTGAAATG 4261

Db 3234 CCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCAGGACCTGAAATG 3293

Qy 4262 ACCCTGTGCTTATTTGAACCTAACCAATCAGTTCGCTTCTCGTTCCTGTTCCGGCGCTTC 4321

Db 3294 ACCCTGTGCTTATTTGAACCTAACCAATCAGTTCGCTTCTCGTTCCTGTTCCGGCGCTTC 3353

Qy 4322 TGTCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGCGGGCCAGTCTCCGAT 4381

Db 3354 TGTCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGCGGGCCAGTCTCCGAT 3413

Qy 4382 AGACTCGCTGCGCCGGGTACCCGTGTTCTCAATAAACCCCTCTTGCAAGTTGCATCCGACTC 4441

Db 3414 TGACTGAGTCGCGCGGGTACCCGTGTTATCCATAAACCCTCTTGCAAGTTGCATCCGACTT 3473

Qy 4442 GTGGTCTCGTGTCTCTGGAGGGTCT-CTCTGAGTGAATGACTACCGTCCAGCGGGT 4500

Db 3474 GTGGTCTCGTGTCTCTGGAGGGTCTCTCTGAGTGATTTGACTACCCGTCCAGCGGGG 3533

Qy 4501 CTTTTCAGTTTCTCCACCTACACAGGTCTCACTAAATCTCTGATGTCGCGGAGGACTC 4560

Db 3534 TCTTTTCATTTGGGGGCTCGTCCGGATCGGGAGACCCCTGCCAGGAGACACCGACCCAC 3593

Qy 4561 CGTACGCCCGGTTTTTGTATTAATAAATAAAGAACAGTGTTCCTTTCAAGCCAGACT 4620

Db 3594 CACCGGGAGTAAAGCTG----- 3610

Qy 4621 ACATCTGACTCTCGGCTTTATAAAGATGTTGAAGGCTCTGTGGACTATCTGCCACA 4680

Db 3611 ----GCTGCTCTCGCGGTTTCGGTGATGACGGTGAACCTCTG-ACATCGAGCTCCC 3665

Qy 4681 CGACTTTTAAAGATTTTATGCTCTCGATGAGGATTTAGTCAATCTATCTCTGCTTA 4740

Db 3666 GGAGACGGTTCACAGCTTGTCTGAACGGATGCCGGAGACAGACCCGTTCAGGGCGC 3725

Qy 4741 TTTTGTGCTTCTCCGTATTTTAAATTTTCTAGTTTTCATCCCTTCTGAGAGCACGGC 4800

Db 3726 GTACGGGGGTGTGGCG----- 3742

Qy 4801 GATTGCAGAGTAGTAAATCTCTGAGGGCAGGCTTCTGTCAAAAGTTGCTGGGCTCAG 4860

Db 3743 ----- 3742

Qy 4861 TGTGAGATTTTGCCATAAAGGGGTCCTGCCCTGTGTACAGACAGATCGGAATCTAGA 4920

Db 3743 ----- 3742

Qy 4921 GTGCATACCTCAGAGTCCCGGGTCTCGATCTCAGGGCATCTTTGCCCTAGAG 4980

Db 3743 -----GGTGTCCGGGGCAGCCATGA----- 3763

Qy 4981 ATCCTCTACCCGCGGACGCATCGTGGCCGGGTACCGAGCTCGAATTCGTAATCATGTCAT 5040

Db 3764 ----- 3763

Qy 5041 AGCTGTTTTCTGTGTGAATTTGTTATCCGCTCACAAATTCACACAAACATACGAGCCGAA 5100

Db 3764 -----CCACGTACGTCAGTCAGCGA 3786

Qy 5101 GCATAAAGTGTAAAGCTGGGGTGCTTAATGAGTGAGCTAACTCATTAATTTGCGTTGC 5160

Db 3787 GTGTATACTGGCTTAACTATGCGGCATCAGACGAGATTGTACTGAGT----- 3835

Qy 5161 GCTCACTGCCCGCTTTCCAGTCGGGAAACCTGCTGTCGAGCTGCATTAATGAATCGGCC 5220

Db 3836 -----GCACCATATGCGGTGTGAATACCGC 3861

Qy 5221 AACCGCGGGGAGAGCGGTTTTCGTATTTGGCGCTCTTCCCGCTTCTCGCTCACTGACT 5280

Db 3862 ACAGATCGGTAAGGAGAAAAATACCGCATCAGCGCTCTTCCGCTTCTCGCTCACTGACT 3921

Qy 5281 CGCTGCGTCTCGTTCGCTGCGCGGAGCGGTATCAGCTCACTCAAAAGCGGTAAATAC 5340

Db 3922 CGCTGCGTCTCGTTCGCTGCGCGGAGCGGTATCAGCTCACTCAAAAGCGGTAAATAC 3981

Qy 5341 GCTTATCCACACAATCAGGGGATACGCAAGGAAAGAACATGTAGCAAAAGCCAGCAAA 5400

Db 3982 GGTATTCACAGAAATCAGGGGATACGCAAGGAAAGAACATGTAGCAAAAGCCAGCAAA 4041

Qy 5401 AGCCAGGAAACCGTAAAAAGCGCGGTTGCTGGCGTTTTTCATAGGCTCCGCCCGCTG 5460

Db 4042 AGCCAGGAAACCGTAAAAAGCGCGGTTGCTGGCGTTTTTCATAGGCTCCGCCCGCTG 4101

Qy 5461 ACAGATCACAAAAAATCGAAGTCAAGTCAAGTGGGAAACCCGACAGGACTATAAA 5520

Db 4102 ACAGATCACAAAAAATCGAAGTCAAGTCAAGTGGGAAACCCGACAGGACTATAAA 4161

Qy 5521 GATACCGCGGTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCTGCGCG 5580

Db 4162 GATACCGCGGTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCTGCGCG 4221

Qy 5581 TTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAAGCTGGCGCTTCTCATAGCTAC 5640

Db 4222 TTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAAGCTGGCGCTTCTCATAGCTAC 4281

Qy 5641 GCTGAGGATCTCATAGTTCGCTGAGTTCGCTCCAGCTGGCTGTGTGTCAGAAC 5700

Db 4282 GCTGAGGATCTCATAGTTCGCTGAGTTCGCTCCAGCTGGCTGTGTGTCAGAAC 4341

Qy 5701 CCCCCTTCAGCCCGACCGCTTATCCGCTTATCCGCTTATCCGCTTATCCGCTTATCCGCT 5760

Db 4342 CCCCCTTCAGCCCGACCGCTTATCCGCTTATCCGCTTATCCGCTTATCCGCTTATCCGCT 4401

Qy 5761 TAAAGACAGACTTATCGCCACTGCGCAGCAGCCACTGGTAAACAGGATTTAGCAGAGCGAGT 5820

Db 4402 TAAAGACAGACTTATCGCCACTGCGCAGCAGCCACTGGTAAACAGGATTTAGCAGAGCGAGT 4461

Qy 5821 ATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGCTTAACTACGGCTACCTAGAGGA 5880

Db 4462 ATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGCTTAACTACGGCTACCTAGAGGA 4521

Qy 5881 CAGTATTTGTTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAGTTGTTAGCT 5940

Db 4522 CAGTATTTGTTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAGTTGTTAGCT 4581

Qy 5941 CTTGATTCGGGCAAAACAAACCCAGCTGGTAGCGGTGTTTTTTTGTTCGAAAGCAGAGA 6000

Db 4582 CTTGATTCGGGCAAAACAAACCCAGCTGGTAGCGGTGTTTTTTTGTTCGAAAGCAGAGA 4641

Qy 6001 TTACGGCGCAAAAAAGGATCTCAAGAGATCTCTTGTGATCTTTTCTACGGGTCTGAGC 6060

Db 4642 TTACGGCGCAAAAAAGGATCTCAAGAGATCTCTTGTGATCTTTTCTACGGGTCTGAGC 4701

Qy	6061	CTCAGTGTGAACGAAAACTCAGGTTAAGGGATTTTGGTCACTGAGATTTATCAAAAAGGATCT	6121
Db	4702	CTCAGTGTGAACGAAAACTCAGGTTAAGGGATTTTGGTCACTGAGATTTATCAAAAAGGATCT	4761
Qy	6121	TCACCTAGATCTCTTTTAAATTTAAAAATGAAAGTTTAAATCAATCTAAAGTATATATCAGT	6180
Db	4762	TCACCTAGATCTCTTTTAAATTTAAAAATGAAAGTTTAAATCAATCTAAAGTATATATCAGT	4821
Qy	6181	AAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAAGCCACCTATCTCAGCGATCTGTC	6240
Db	4822	AAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAAGCCACCTATCTCAGCGATCTGTC	4881
Qy	6241	TATTTTGGTTCATTCATAGTTGCTGTGATCTCCCGTCTGTGTAGATAAATCAGATACGGAGG	6300
Db	4882	TATTTTGGTTCATTCATAGTTGCTGTGATCTCCCGTCTGTGTAGATAAATCAGATACGGAGG	4941
Qy	6301	GCCTTACCATCTGGCCCCAGTGTCTCAATGATACCGCAGAGCCCAACGCTCACCGGCTCCAG	6360
Db	4942	GCCTTACCATCTGGCCCCAGTGTCTCAATGATACCGCAGAGCCCAACGCTCACCGGCTCCAG	5001
Qy	6361	ATTTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAAGTGGTCTCTGCAACTT	6420
Db	5002	ATTTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAAGTGGTCTCTGCAACTT	5061
Qy	6421	TATCCGGCTCCATCCAGTCTATTAATTTGTTGTCGGGGAAGCTAGAGTAAAGTAGTTTCGCCAG	6480
Db	5062	TATCCGGCTCCATCCAGTCTATTAATTTGTTGTCGGGGAAGCTAGAGTAAAGTAGTTTCGCCAG	5121
Qy	6481	TTAATAGTTTTCGCAACGTTTGGCAATGCTACAGGC - TCGTGGTCTCAGCTCGTCTG	6539
Db	5122	TTAATAGTTTTCGCAACGTTTGGCCATTTGCTCAGGCAATCGTGGTGTCAAGCTCGTCTG	5181
Qy	6540	TTGGTATGGCTTCATTCAGCTCCGGTTCCCAACCATCAAGCGCAGTTTACATGATCCGCCCA	6599
Db	5182	TTGGTATGGCTTCATTCAGCTCCGGTTCCCAACCATCAAGCGCAGTTTACATGATCCGCCCA	5241
Qy	6600	TGTTGTCAAAAAAGCGGTTAGTCTCTTCGGTCTCTCCGATCGTTTGTGAGAAGTAAAGTTGG	6659
Db	5242	TGTTGTCAAAAAAGCGGTTAGTCTCTTCGGTCTCTCCGATCGTTTGTGAGAAGTAAAGTTGG	5301
Qy	6660	CCGCGAGTTTATCACTCATGGTTATGGCAGCAGCTGCATATTTCTTCTTACTGTCTATGCCAT	6719
Db	5302	CCGCGAGTTTATCACTCATGGTTATGGCAGCAGCTGCATATTTCTTCTTACTGTCTATGCCAT	5361
Qy	6720	CCGTAAAGATGCTTTTCTGTGACTGGTGAGTACTCAAACCAAGTCAATCTCTGAGAATAGTGTA	6779
Db	5362	CCGTAAAGATGCTTTTCTGTGACTGGTGAGTACTCAAACCAAGTCAATCTCTGAGAATAGTGTA	5421
Qy	6780	TGCGCGACCGAGTTGCTCTTTGCCGCGTCAATACGGGATAATACCGCGCCCAATAGCA	6839
Db	5422	TGCGCGACCGAGTTGCTCTTTGCCGCGCGTCAACACGGGATAATACCGCGCCCAATAGCA	5481
Qy	6840	GAACTTTTAAAGTCTCATCATTTGGAANAACGTTCTTTCGGGGCGAAAACTCTCAAGGATCT	6899
Db	5482	GAACTTTTAAAGTCTCATCATTTGGAANAACGTTCTTTCGGGGCGAAAACTCTCAAGGATCT	5541
Qy	6900	TACCGCTGTTGAGATTCAGTTTCGATGTAAACCCACTCGTGCACCCAACTGATCTTTCAGCAT	6959
Db	5542	TACCGCTGTTGAGATTCAGTTTCGATGTAAACCCACTCGTGCACCCAACTGATCTTTCAGCAT	5601
Qy	6960	CTTTTACTTTTACCAGCGTTTCTCGGTTGAGCAAAAAACGGAAGGCAAAATGCCGCAAAAA	7019
Db	5602	CTTTTACTTTTACCAGCGTTTCTCGGTTGAGCAAAAAACGGAAGGCAAAATGCCGCAAAAA	5661
Qy	7020	AGGGAATAAGGGCGACACGGAAATGTTGAATACATCTCATCTCTCTCTTTTCAATATTATT	7079
Db	5662	AGGGAATAAGGGCGACACGGAAATGTTGAATACATCTCATCTCTCTCTTTTCAATATTATT	5721
Qy	7080	GAAACATTTATCAGGGTTATTGTCTCATGACGGGATACATATTTGAAATGATTTTAGAAAA	7139
Db	5722	GAAACATTTATCAGGGTTATTGTCTCATGACGGGATACATATTTGAAATGATTTTAGAAAA	5781
Qy	7140	ATAAACAAATAGGGGTTTCGCGCACATTTTCCCGGAAAAAGTGCCACCTGACGTCTTAAGAAA	7199

Db	5782	ATAAACAAATAGGGGTTCCGGCGCAATTTCCCGAANAAGTGCCACCTGACGTCTTAAGAAA	5841
Qy	7200	CCATTATTATCATGACATTAACTCTATAAANAATAGCGGTATCACGAGGCCCTTTTCGTCT	7257
Db	5842	CCATTATTAATGACATTAACTCTATAAANAATAGCGGTATCACGAGGCCCTTTTCGTCT	5899
RESULT 6			
LOCUS	AX491314	5782 bp	DNA linear PAT 16-AUG-2002
DEFINITION	Sequence 1 from Patent WO0234929.		
ACCESSION	AX491314		
VERSION	AX491314.1	GI:22324009	
KEYWORDS			
SOURCE	synthetic construct		
ORGANISM	other sequences; artificial sequences.		
REFERENCE	1		
AUTHORS	Kohn, E.Y. and Daley, G.O.		
TITLE	Expression vectors and uses thereof		
JOURNAL	Patent: WO 0234929-A 1 02-MAY-2002;		
	WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US)		
FEATURES	Location/Qualifiers		
source	1..5782		
	/organism="synthetic construct"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32630"		
ORIGIN	/note="Synthetically generated nucleic acid"		
Query Match	42.0%;	Score 3213;	DB 6; Length 5782;
Best Local Similarity	73.3%;	Pred. No. 0;	
Matches 4860;	Conservative 0;	Mismatches 905;	Indels 864; Gaps 24;
Qy	552	TGAAAGACCCACCTGTAGTTTGGCAAGCTAGCTTAAGTAAACGCCAATTTTGCAGGCGAT	611
Db	3	TGAAAGACCCACCTGTAGTTTGGCAAGCTAGCTTAAGTAAACGCCAATTTTGCAGGCGAT	62
Qy	612	GG-AAAATACATACTCAGAAATCAGAAAGTTTCAGATCAAGGTTAGGAACAGA-GAGACAG	669
Db	63	GGAANAATACATACTCAGAAATGAAAGTTTCAGATCAAGGTCAGGAACAGATGGAACAG	122
Qy	670	CAGAAATATGGGCCAAACAGGATATCTGTGGTAAACAGTTCCTGCCCC-GCTCAGGGCCAA	728
Db	123	CTGAATATGGSCCAAAGCGGATATCTGTGGTAAACAGTTCCTGCCCCGGCTCAGGGCCAA	182
Qy	729	GAAAGTTGGAACAGGAATATATGGGCCAAACAGGATATCTGTGGTAAACAGTTCCTGCC	788
Db	183	GAAAGATGGAACAGCTGAAATATGGGCCAAACAGGATATCTGTGGTAAACAGTTCCTGCC	242
Qy	789	CCGGCTCAGGGCCAAAGAACAGATGTCCTCCAGATGCGGTCCCGCCCTCAGCAGGTTCTAG	848
Db	243	CCGGCTCAGGGCCAAAGAACAGATGTCCTCCAGATGCGGTCCCGCCCTCAGCAGGTTCTAG	302
Qy	849	AGAAACCATCAGATGTTTCCAGGGTGCCTCCAGGACCTGAAATGACCTTCGCTTATTTG	908
Db	303	AGAACCATCAGATGTTTCCAGGGTGCCTCCAGGACCTGAAATGACCTTCGCTTATTTG	362
Qy	909	AATTAACCAATCAGTTTCGCTTCTCGTTTCGTTTCGGGGCTTCCTGCTCCCGAGCTCAAT	968
Db	363	AATTAACCAATCAGTTTCGCTTCTCGTTTCGTTTCGGGGCTTCCTGCTCCCGAGCTCAAT	422
Qy	969	AAAGAGCCCAACACCTCCTCAGCGCGCCAGTCTCCGATAGACTGCGTCCCGCCGG	1028
Db	423	AAAGAGCCCAACACCTCCTCAGCGCGCCAGTCTCCGATAGACTGCGTCCCGCCGG	482
Qy	1029	TACCGGTATTCCTCAATTAAGCTCTTGTGTTTCATCCGAACTCGTGACCTCGCTGATCC	1088
Db	483	TACCGGTATTCCTCAATTAAGCTCTTGTGTTTCATCCGAACTCGTGACCTCGCTGATCC	541
Qy	1089	TTGGAGGGTCTCTCAGATTGATTGATGCCCC-ACCTCGGGGGTCTTTTCATTTTGGAGGT	1147



3307 GGGGGGGCGAGTGCACACGAGGGGGCTGGACTTCGCCCTGGATCCAAACTCTGCTAC 3366  
Db |||||  
2675 TTTGGCGCTTAAATTTGAGCTCGAGCGCATATCTACGCGCGGA-----GAAACCGCCTC 2730  
Qy |||||  
3367 CTGCTGGATGAAATCCTCTTCATCTATGGTGTCACTTCTCACTGCCCTTGTCTCTGAGATG 3426  
Db |||||  
2731 GCGGTGATGTTGCTGCGCTGGAGTGACGGGAGTTATCT----- 2768  
Qy |||||  
3427 AAGTTACAGAGCGCAGAGCCCGCCCGCTACACAGCAGGGCCAGAACCACTCTATTAAC 3486  
Db |||||  
2769 -----TGAAGATCMAGATATGTGGCGGATGAGCGGATTCGAGCGCAAAAC 2814  
Qy |||||  
3487 GAGCTCAATCTAGGACGAAGAGAGTACGATGTTTTGGACAGAGAGTGGCGGGAC 3546  
Db |||||  
2815 GGTCTGCGCTCGGGACGCGGAATTGAATTTATGGCCCAACACAGAGTGGCGCGGGCGAC 2874  
Qy |||||  
3547 CCTGAGATGGGGGAAGCCGAGAGGAAGAACCTCAGGAAGGCCCTGACAAATGAATCTG 3606  
Db |||||  
2875 TTCCAGTT-----CAAATCAGCGCTACAGTCAACAGCACTGATGGAAACAGCCAT 2928  
Qy |||||  
3607 CAGAAAGATAAGATGGCGGAGCCCTACAGTGAGATTTGGGNTGAAAGCGAGCGCGGAGG 3666  
Db |||||  
2929 CGCCATCTGCTGCACGCGGAAGAACCGACATGGCTGTTATACGACGGTTTCCATATGGG 2988  
Qy |||||  
3667 GGCAAGGGCAGATGGCTTTACAGGGTCTCAGTACAGCCACCAAGGACACCTACGAC 3726  
Db |||||  
2989 ATTGGTGGCAGCACTCTCGAGGCCGTCAGTATGCGCGGAATTCAGCTGAGCGCCGGT 3048  
Qy |||||  
3727 GCCCTTACATGACAGGCCCTCGCCCTCGCTAACTCGACGCGCGCGGATCCGGATTAG 3786  
Db |||||  
3049 CGCTACCATTTACAGTTGGTCTGGTGTCAAAATAAATAAATACCGGGCAGGCCATGTCTG 3108  
Qy |||||  
3787 TCCAATTTGTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTGTACTCAACAATATC 3846  
Db |||||  
3109 CCGCTATTTCGTAAGGAATCCATATGT-----ACTATTGA 3147  
Qy |||||  
3847 ACCAGCTGAAGCCTATAGAGTACGAGCCATAGATAAATAAAGATTTATTTAGTCTCC 3906  
Db |||||  
3148 AACTCGAGCGCGCCGACAGCAGTGGTCGAGGATAAATAAAGATTTATTTAGTCTCC 3207  
Qy |||||  
3907 AGAAAAGGGGGAATGAAGACCCACCTGTAGTTTGGCAAGCTAGCTTAAGTAAGC 3966  
Db |||||  
3208 AGAAAAGGGGGAATGAAGACCCACCTGTAGTTTGGCAAGCTAGCTTAAGTAAGC 3267  
Qy |||||  
3967 CATTTTGAAGGCATGG-AAAATACATACTGAGATAGAGAAGTTCCAGATCAAGGTTAG 4025  
Db |||||  
3268 CATTTTGAAGGCATGGAAAATAACATACTGAGATAGAGAAGTTCCAGATCAAGTCAG 3327  
Qy |||||  
4026 GAACAGA-GAGACAGCAGATATGGGCCAAACAGAGATCTGTGTAAGCAGTTCCTGCC 4084  
Db |||||  
3328 GAACAGATGGAAACAGCTGAATATGGGCCAAACAGGATATCTGTGTAAGCAGTTCCTGCC 3387  
Qy |||||  
4085 CC-GCTCAGGCGCAGACAGTTGGACAGAGATATGGGCCAAACAGGATATCTGTGG 4143  
Db |||||  
3388 CCGGCTCAGGCGCAGAAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGG 3447  
Qy |||||  
4144 TAAGCAGTTCTGTGCCCGGCTCAGGGCCAAAGAACAGATGGTCCCAAGATGGGTCCTGCC 4203  
Db |||||  
3448 TAAGCAGTTCTGTGCCCGGCTCAGGGCCAAAGAACAGATGGTCCCAAGATGGGTCCTGCC 3507  
Qy |||||  
4204 CTCAGCAGTTCTTAGAGAACCATCAGATGTTTTCAAGGGTGGCCCAAGGACCTGAAATGAC 4263  
Db |||||  
3508 CTCAGCAGTTCTTAGAGAACCATCAGATGTTTTCAAGGGTGGCCCAAGGACCTGAAATGAC 3567  
Qy |||||  
4264 CCTGTGCTTATTTGAATACCAATCAGTTGCTTCTCGCTTCTGTGTCGGCGCTTCTG 4323  
Db |||||  
3568 CCTGTGCTTATTTGAATACCAATCAGTTGCTTCTCGCTTCTGTGTCGGCGCTTCTG 3627  
Qy |||||  
4324 CTCGCCGAGCTCAATAAAGAGCCCAACAAACCTCAGTCTGGCGCGCAGTCTCTCGGATAG 4383  
Db |||||  
3628 CTCGCCGAGCTCAATAAAGAGCCCAACAAACCTCAGTCTGGGGGCCAGTCTCTCGATG 3687  
Qy |||||  
4384 ACTGGTGGCGGGTACCGGTGTTCTCAATAAACCCTCTTTGACAGTTGCATCCGACTCGT 4443

3688 ACTGAGTCGCCCGGTACCCGTGTATCCAATAAAACCCCTCTTGCAGTTGCATCCGACTTGT 3747  
Qy |||||  
4444 GGTCTCGCTGTTCTTGGAGGGTCT-CTCTGAGTGATTGACTACCCCGTACGCGGGTCT 4502  
Db |||||  
3748 GGTCTCGCTGTTCTTGGAGGGTCTCTCTGAGTGATTGACTACCCCGTACGCGGG--- 3804  
Qy |||||  
4503 TTCAGTTTCTCCCACTACACAGGTCTCACTAAACATTCCTGATGTGCGCAGGGACTCG 4562  
Db |||||  
3805 ----- 3804  
Qy |||||  
4563 TCAGCCCGTTTTTGTATTATAATAAATGCAAGACAGTGTTCCTTCAAGCCAGACTAC 4622  
Db |||||  
3805 ----- 3804  
Qy |||||  
4623 ATCTGACTCTCGGCTTTATAAAAGATGTTGAAGGGCTCTGTGAGACTATCTGCCACACG 4682  
Db |||||  
3805 ----- 3804  
Qy |||||  
4683 ACTTTTTAAGATTTTTATGCTCTGGATGAGGATTTAGTCAATCTATCCTGCTTAT 4742  
Db |||||  
3805 ----- 3804  
Qy |||||  
4743 TTGCTGGCTTCTCGTATTTTAAATTTCTAGTTTGCACCTCCCTTCTGAGAGCAGCGGA 4802  
Db |||||  
3805 ----- 3804  
Qy |||||  
4803 TTGCAGAGTAGTAAATCTCTGAGGGCAGGCTTCTGTGAAAAGTTGCTTGGGCTCAGTG 4862  
Db |||||  
3805 ----- 3804  
Qy |||||  
4863 TGAGATTTTGCCATATAAAAGGGTCTGCCCCCTGTGTACAGACAGATCGSAATCTAGAGT 4922  
Db |||||  
3805 ----- 3804  
Qy |||||  
4923 GCATACTCAGAGTCCCGCGGTTCCGGGGTCTGTATCTCAGGGCATCTTTGCCGTAGAGAT 4982  
Db |||||  
3805 ----- 3804  
Qy |||||  
4983 CCTCTACCGCGAGCGCATCGTGGCGGGTACCGAGTCCGAATTCGTAATCATGTGTCATAG 5042  
Db |||||  
3805 ----- 3804  
Qy |||||  
5043 CTGTTTCTGTGTGAAATTTGTTATPCGCTCAAAATCCACACAATACAGAGCCGGAAGC 5102  
Db |||||  
3805 ----- 3804  
Qy |||||  
5103 ATAAAGTGAAGCCTGGGGTGCCCTAATGAGTGAGTAACTCACATTAATTTGCGTTGCGC 5162  
Db |||||  
3805 ----- 3804  
Qy |||||  
5163 TCACCTGCCCGTTTCCAGTCCGGAAACCTGTGTCGAGCTGCATTAATGAATCGGCCAA 5222  
Db |||||  
3805 -----GTCCTTCATTCGCAATTAATGAATCGGCCAA 3835  
Qy |||||  
5223 CGCGCGGGAGAGCGGTTTGGTATTTGGGGCTCTTCCGCTTCTCGCTCACTGACTCG 5282  
Db |||||  
3836 CGCGCGGGAGAGCGGTTTGGTATTTGGGGCTCTTCCGCTTCTCGCTCACTGACTCG 3895  
Qy |||||  
5283 CTGGCTCGGTCTGTTGGCTGCGCGAGCGGTATCAGTCTCACTAAAGCGGTTAATACGG 5342  
Db |||||  
3896 CTGGCTCGGTCTGTTGGCTGCGCGAGCGGTATCAGTCTCACTCAAAAGCGGTTAATACGG 3955  
Qy |||||  
5343 TTATCCACAGAACTCAGGGGATACCGAGGAAGAACATGTGAGCAAAAGGCCAGCAAAAG 5402  
Db |||||  
3956 TTATCCACAGAACTCAGGGGATACCGAGGAAGAACATGTGAGCAAAAGGCCAGCAAAAG 4015  
Qy |||||  
5403 GCCAGGAACCTGAAAAAGGCGCGTGTGTCGCGTTTTTCCATAGGCTCCGCCCCCTGAC 5462  
Db |||||  
4016 GCCAGGAACCTGAAAAAGGCGCGTGTGTCGCGTTTTTCCATAGGCTCCGCCCCCTGAC 4075  
Qy |||||  
5463 GAGCATCACAAAAATCCAGCTCAAGTCAAGGTGGCGAAACCCGACAGGACTATAAAGA 5522  
Db |||||



Db 4076 GAGCATACAAAAATCGACGCTCAAGTCAGAGGTGGCAAAACCCGACAGGACTATAAAGA 4135  
Qy 5523 TACACGGCTTCCCTCGAGCTCCCTCGTGGCTCTCCCTGTTCCGACCCCTGCGCTT 5582  
Db 4136 TACACGGCTTCCCTCGAGCTCCCTCGTGGCTCTCCCTGTTCCGACCCCTGCGCTT 4195  
Qy 5583 ACCGATACCTGTCGCCCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGC 5642  
Db 4196 ACCGATACCTGTCGCCCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGC 4255  
Qy 5643 TGTAGTATCTCAGTTTGGGTGATGTTGCTTCCCTCAAGCTGGGCTGTGTGACGAAACC 5702  
Db 4256 TGTAGTATCTCAGTTTGGGTGATGTTGCTTCCCTCAAGCTGGGCTGTGTGACGAAACC 4315  
Qy 5703 CCGTTTCAGCCGACCGCTTTCGCGCTTATCGGTTAACTATCGTTCGAGTCCACCCGTA 5762  
Db 4316 CCGTTTCAGCCGACCGCTTTCGCGCTTATCGGTTAACTATCGTTCGAGTCCACCCGTA 4375  
Qy 5763 AGACAGCACTTATCGGCACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAGGTAT 5822  
Db 4376 AGACAGCACTTATCGGCACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAGGTAT 4435  
Qy 5823 GTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTTAACTACGGCTACACTAGAGGACA 5882  
Db 4436 GTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTTAACTACGGCTACACTAGAGGACA 4495  
Qy 5883 GTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTTACCTTCGGAAGAGTTGGTAGCTCT 5942  
Db 4496 GTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTTACCTTCGGAAGAGTTGGTAGCTCT 4555  
Qy 5943 TGATCCGGCAAAACAAACACCGCTGCTAGCGGTGGTGTGTTTGTGTTGCAAGCAGAGATT 6002  
Db 4556 TGATCCGGCAAAACAAACACCGCTGCTAGCGGTGGTGTGTTTGTGTTGCAAGCAGAGATT 4615  
Qy 6003 ACGCGCAGAAAAAAGATCTCAAGAAAGTCTTTTGAATCTTTTCTACGGGGTCTGACGT 6062  
Db 4616 ACGCGCAGAAAAAAGATCTCAAGAAAGTCTTTTGAATCTTTTCTACGGGGTCTGACGT 4675  
Qy 6063 CAGTGAACGMAAACHTCAGTTAAGGATTTGGTTCATGAGATTATCAAAAGGATCTTC 6122  
Db 4676 CAGTGAACGMAAACHTCAGTTAAGGATTTGGTTCATGAGATTATCAAAAGGATCTTC 4735  
Qy 6123 ACTAGATCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATGAGTAA 6182  
Db 4736 ACTAGATCTTTT-----GGCGCGCGCCAAATCAATCTAAAGTATATATGAGTAA 4788  
Qy 6183 ACTTGTCTGACAGTTACAAATGCTTAATCAGTGAAGCCACCTATCTCAGCGATCTCTTA 6242  
Db 4789 ACTTGTCTGACAGTTACAAATGCTTAATCAGTGAAGCCACCTATCTCAGCGATCTCTTA 4848  
Qy 6243 TTTCTGTTCCATAGTTGCTGACTCCCGTGTGTAGATAACTACGATACGGAGGGC 6302  
Db 4849 TTTCTGTTCCATAGTTGCTGACTCCCGTGTGTAGATAACTACGATACGGAGGGC 4908  
Qy 6303 TTACCATCTGGCCAGTGTGCAATGATACCGGAGACCCAGCTCACCGGCTCCAGAT 6362  
Db 4909 TTACCATCTGGCCAGTGTGCAATGATACCGGAGACCCAGCTCACCGGCTCCAGAT 4968  
Qy 6363 TTATCAGCAATAAACACGACCGCGGAAGGGCGGAGAGTGGTCTCTGCAACTTTA 6422  
Db 4969 TTATCAGCAATAAACACGACCGCGGAAGGGCGGAGAGTGGTCTCTGCAACTTTA 5028  
Qy 6423 TCCGCTCCATCCAGTCTATTATGTTGCGGGAAGCTAGAGTAGTAGTTCCGCGATT 6482  
Db 5029 TCCGCTCCATCCAGTCTATTATGTTGCGGGAAGCTAGAGTAGTAGTTCCGCGATT 5088  
Qy 6483 AATAGTTGGCAACGTTGTTGCAATGCTACAGGC-TGTTGTTGTCAGCTCGTGGTT 6541  
Db 5089 AATAGTTGGCAACGTTGTTGCAATGCTACAGGCATGTTGTTGTCAGCTCGTGGTT 5148  
Qy 6542 GGTATGGCTTCAATTCAGCTCCGTTCCCAACGATCAAGCGGAGTTACATGATCCCCCATG 6601  
Db 5149 GGTATGGCTTCAATTCAGCTCCGTTCCCAACGATCAAGCGGAGTTACATGATCCCCCATG 5208

Qy 6602 TTGTGCAAAAAAGCGTTAGTCTTCCGATCCTCGATCGTTGTGTCAGAGTAAGTTGGCC 6661  
Db 5209 TTGTGCAAAAAAGCGTTAGTCTTCCGATCCTCGATCGTTGTGTCAGAGTAAGTTGGCC 5268  
Qy 6662 CGAGTGTATCACTCATGTTATGCGACACTGCAATAATCTTCTTACTGTCATGCAATCC 6721  
Db 5269 CGAGTGTATCACTCATGTTATGCGACACTGCAATAATCTTCTTACTGTCATGCAATCC 5328  
Qy 6722 GTAAGATGCTTTTCTGTGACTGCTGAGTACTCAACCAAGTCAATCTGAGAAATAGTGTG 6781  
Db 5329 GTAAGATGCTTTTCTGTGACTGCTGAGTACTCAACCAAGTCAATCTGAGAAATAGTGTG 5388  
Qy 6782 CGCGACCGAGTGTCTTTCGCGCGCTCAATACCGGATAATACCGCGCCACATAGCAGA 6841  
Db 5389 CGCGACCGAGTGTCTTTCGCGCGCTCAATACCGGATAATACCGCGCCACATAGCAGA 5448  
Qy 6842 ACTTTAAAGTCTCATCATTTGAAAAAGTCTTCTCGGCGGAAAACTCTCAAGGATCTTA 6901  
Db 5449 ACTTTAAAGTCTCATCATTTGAAAAAGTCTTCTCGGCGGAAAACTCTCAAGGATCTTA 5508  
Qy 6902 CGCTGTTGAGATCAGTTTGAATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCT 5961  
Db 5509 CGCTGTTGAGATCAGTTTGAATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCT 5568  
Qy 6962 TTTTACTTTTACCAGCGTTTCTCGGTGAGCAAAAAAGCAAAATGCCCCCAAAAAAG 7021  
Db 5569 TTTTACTTTTACCAGCGTTTCTCGGTGAGCAAAAAAGCAAAATGCCCCCAAAAAAG 5628  
Qy 7022 GGAATAAGCGGACACGGAATGTTGAATCTCATACTCTTCTTTTCAATATTATGA 7081  
Db 5629 GGAATAAGCGGACACGGAATGTTGAATCTCATACTCTTCTTTTCAATATTATGA 5688  
Qy 7082 AGCATTTATCAGGTTTATGTTCTCATGAGCGATACATATTTGAATGTTTGAATAAT 7141  
Db 5689 AGCATTTATCAGGTTTATGTTCTCATGAGCGATACATATTTGAATGTTTGAATAAT 5748  
Qy 7142 AAACAAATAGGGGTTCCGCGCACATTTCC 7170  
Db 5749 AAACAAATAGGGGTTCCGCGCACATTTCC 5777

## RESULT 7

E23357 7372 bp DNA linear PAT 18-JUN-2001  
LOCUS Virus vector system expressing apoptosis-related gene.  
DEFINITION E23357  
ACCESSION E23357  
VERSION E23357.1 GI:13024380  
KEYWORDS JP 1999075859-A/3.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 7372)  
AUTHORS Hirofumi,H.  
TITLE Virus vector system expressing apoptosis-related gene  
JOURNAL Patent: JP 1999075859-A 3 23-MAR-1999;  
R B R JENSERU KK  
COMMENT OS Homo sapiens (human)  
PN JP 1999075859-A/3  
PD 23-MAR-1999  
PF 08-SEP-1997 JP 1997259235  
PR  
PI HIROFUMI HAMADA  
PC C12N15/09,C12N5/10,C12N7/00//A61K35/76,A61K48/00,(C12N5/10,PC  
C12R1:91),  
PC (C12N7/00,C12R1:92),C12N15/00,C12N5/00,(C12N5/00,C12R1:91) CC  
Strandedness: Double;  
CC Topology: Linear;  
FH Key Location/Qualifiers  
FT source 1..7372 /organism='Homo sapiens (human)'.  
FT Location/Qualifiers



source	1..7372	/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
ORIGIN			
Query Match	41.6%;	Score 3183.2;	DB 6; Length 7372;
Best Local Similarity	70.7%;	Pred. No. 0;	
Matches 4893; Conservative	0;	Mismatches 1503;	Indels 528; Gaps 28;
Qy	843	TTCTAGAGAACCATCAGATGTTTCCAGGGTGCCCAAGGACCTGAATGACCCCTGTGCCT	902
Db	467	TGCTAGAGAACCATCAGATGTTTCCAGGGTGCCCAAGGACCTGAATGACCCCTGTGCCT	526
Qy	903	TATTGAACTAACCAATCAGTTCGCTTCTCGCTTCTGTTGCGCGCTCTCTGCCCGAG	962
Db	527	TATTGAACTAACCAATCAGTTCGCTTCTCGCTTCTGTTGCGCGCTCTCTGCCCGAG	586
Qy	963	CTCAATAAAGAGCCCAACCCCTCACTCGGCGGCCAGTCTCCGATAGACTGCGTGC	1022
Db	587	CTCAATAAAGAGCCCAACCCCTCACTCGGCGGCCAGTCTCCGATAGACTGCGTGC	646
Qy	1023	CCCGGTACCCGCTATTTCCCAATAAAGCCTCTTGCTGTTTGCATCCGAATCCTGACTCGC	1082
Db	647	CCCGGTACCCGCTATTTCCCAATAAAGCCTCTTGCTGTTTGCATCCGAATCCTGACTCGC	705
Qy	1083	TGATCCTTTGGGAGGGTCTCTCAGATTGATTGACTGCCCC-ACCTCGGGGGTCTTTTCATTT	1141
Db	706	TGTTCCCTTTGGGAGGGTCTCTCTGAGTGAATTGACTACCCGTCAGCGGGGGTCTTTTCATTT	765
Qy	1142	GGAGGTTCCACCGAGATTGGAGACCCCTGCCCAAGGACCAACCGACCCCCCGCGGGAG	1201
Db	766	GGGGGCTCGTCCGGGATCGGAGACCCCTGCCCAAGGACCAACCGA-CCCAACACCGGGAG	824
Qy	1202	GTAAGCTGGCCAGCAACTTATCTGTCTGTCTGTCGAGTTGTCTAGTCTCTACACTGATTTT	1261
Db	825	GTAAGCTGGCCAGCAACTTATCTGTCTGTCTGTCGAGTTGTCTAGTCTCTACACTGATTTT	884
Qy	1262	ATGGCCCTCGCTCGGTACTAGTTAGCTAACTAGCTCTGTACTGCGGACCCGCTGGTGA	1321
Db	885	ATGGCCCTCGCTCGGTACTAGTTAGCTAACTAGCTCTGTACTGCGGACCCGCTGGTGA	944
Qy	1322	ACTGACGAGTTCCGAACACCCGGCGCAACCCCTGGGAGACGTCCACAGGACTTCGGGGGC	1381
Db	945	ACTGACGAGTTCCGAACACCCGGCGCAACCCCTGGGAGACGTCCACAGGACTTCGGGGGC	1004
Qy	1382	CGTTTTTGTGGCCGACTGAGTCTTAATAATCCGATCGTTTGAAGCTCTTTGGTGCACC	1441
Db	1005	CGTTTTTGTGGCCGACTGAGTCTTAATAATCCGATCGTTTGAAGCTCTTTGGTGCACC	1064
Qy	1442	CCCTTAGAGAGGGATATGTGGTTCTGGTAGGAGACGAGAACCTTAAACAGTTTCCCGCC	1501
Db	1065	CCCTTAGAGAGGGATATGTGGTTCTGGTAGGAGACGAGAACCTTAAACAGTTTCCCGCC	1124
Qy	1502	TCGCTCTGAATTTTGGCTTTCCGTTTGGGACCGAAGCCGCGCGCGCTCTGTCTGCTG	1561
Db	1125	TCGCTCTGAATTTTGGCTTTCCGTTTGGGACCGAAGCCGCGCGCGCTCTGTCTGCTG	1184
Qy	1562	CAGCATCGTCTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT	1621
Db	1185	CAGCATCGTCTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT	1244
Qy	1622	CCCGGGCTAGACTGTTTACCACTCCCTTAAGTTTGAACCTTTAGTCTCACTGGAAGATGTCGA	1681
Db	1245	CC-----AGACTGTTTACCACTCCCTTAAGTTTGAACCTTTAGTCTCACTGGAAGATGTCGA	1298
Qy	1682	CGCGATCGCTCAACCAAGTCGGTAGATGTCAAGAAGAGAGCTTGGGTACTTCTGCTC	1741
Db	1299	CGCGATCGCTCAACCAAGTCGGTAGATGTCAAGAAGAGAGCTTGGGTACTTCTGCTC	1358
Qy	1742	TGCGAATGGCCAACTTTTAAAGTCGGATGCGCGGAGACCGGACCTTTTAAACCGAGACT	1801
Db	1359	TGCGAATGGCCAACTTTTAAAGTCGGATGCGCGGAGACCGGACCTTTTAAACCGAGACT	1418
Qy	1802	CATCACCCAGGTTAAGATCAAGGTCCTTTTACCTGGCCCGCATGGACACCAGACCAAGT	1861
Db	1419	CATCACCCAGGTTAAGATCAAGGTCCTTTTACCTGGCCCGCATGGACACCAGACCAAGT	1478
Qy	1862	CCCTTACATCGTGACCTGGGAGGCTTGGCTTTTGAACCCCTCCCTGGGTCAAGCCCTT	1921
Db	1479	CCCTTACATCGTGACCTGGGAGGCTTGGCTTTTGAACCCCTCCCTGGGTCAAGCCCTT	1538
Qy	1922	TGTACACCTTAAGCCTCGCCTCTCTTCTTCCATCGCCCGCCCTCTCTCCCTTTGAAC	1981
Db	1539	TGTACACCTTAAGCCTCGCCTCTCTTCTTCCATCGCCCGCCCTCTCTCCCTTTGAAC	1598
Qy	1982	TCCTGTTTGAACCCCGCTCGATCCTCCCTTTATCCAGCCCTCACTCTCTCTAGGGCG	2041
Db	1599	TCCTGTTTGAACCCCGCTCGATCCTCCCTTTATCCAGCCCTCACTCTCTCTAGGGCG	1658
Qy	2042	CCCATATGCGCATATGAGATCTTATATGGGGGACCCCGCCCTTGTAAACTTCCCTGA	2101
Db	1659	-----CCATATGAGATCTTATATGGGGGACCCCGCCCTTGTAAACTTCCCTGA	1708
Qy	2102	CCCTGACATGACAAAGATTACTAACAGCCCTCTCTCCAAAGCTCACTTACAGGCT-TCTA	2160
Db	1709	CCCTGACATGACAAAGATTACTAACAGCCCTCTCTCCAAAGCTCACTTACAGGCTCTCTA	1768
Qy	2161	CTTAGTCCAGCACGAAGTCTGGAGACCTCTGGCGGACGCTTACCAAGAACCACTGGACCG	2220
Db	1769	CTTAGTCCAGCACGAAGTCTGGAGACCTCTGGCGGACGCTTACCAAGAACCACTGGACCG	1828
Qy	2221	ACCGTGGTACCTCACCCCTTACCGAGTGGGGACACAGTGTGGGTCCGCGACACAGAC	2280
Db	1829	ACCGTGGTACCTCACCCCTTACCGAGTGGGGACACAGTGTGGGTCCGCGACACAGAC	1888
Qy	2281	TAAAGACTAGAACCTCGCTGGAAGAGACTTACACAGTCTCTGTGACACACCCACACGC	2340
Db	1889	TAAAGACTAGAACCTCGCTGGAAGAGACTTACACAGTCTCTGTGACACACCCACACGC	1948
Qy	2341	CCTCAAGTAGACGGCATCGCAGCTTGGATACACCGCGCCCAACGCTGAAGGCTGCGACCC	2400
Db	1949	CCTCAAGTAGACGGCATCGCAGCTTGGATACACCGCGCCCAACGCTGAAGGCTGCGACCC	2008
Qy	2401	CGGGGGTGACCATCTCTTAGACTGCCATGGGATGGAGCTGTATCATCTCTCTTCTTGTA	2460
Db	2009	CGGGGGTGACCATCTCTTAGACTGCCGATCCCAAGTGTGGTGTGGTGTGGTGTGGTGTGGT	2068
Qy	2461	GCAACAGCTACAG-----	2473
Db	2069	AAAAATGTCTCAGAGCAACCGGGAGCTGGTGTGACTTTCTCTCTCAAGCTTTTCCCA	2128
Qy	2474	-----	2473
Db	2129	GAAAGGATACAGCTGGAGTCAGTTTAGTGTGTAAGAGAGAACAGGACTGAGGCCCCAGA	2188
Qy	2474	-----GTGTCCA	2480
Db	2189	AGGACTGAATCGGAGATGGAGACCCCGAGTCCCATCAATGGCAACCCATCTCTGGCACCT	2248
Qy	2481	CTCCGACATCCAGCTGACCCAGAGCCCAAGCAGCTGAGCGCCAGCGTGGGTGACAGAGT	2540
Db	2249	GGCAGACAGCCCCCGGTGATGGAGCCACTGCGCACAGCAGCAGTGTGGATGCCCGGA	2308
Qy	2541	GACCATCACCTGTAAAGC-----CAGTCAAGGATGTGGTACT	2577
Db	2309	GGTATCCCCATGGCAGCATGAAGCAAGCGCTGAGGGAGCGGAGCGAGTTTGAAGT	2368
Qy	2578	TCTGTAGCTTGGTACAGCAGAGACCCAGGTA---AGGCTCAAAAGCTGCTGATCTTCTGG	2634
Db	2369	CGGGTACCGGCGGCAATTCAGTGACCTTGAATCCAGCTCCACATCACCCAGGAGACAGC	2428
Qy	2635	ACATCCACCCGGCACACTGGTGTGCCAAGCAGATTACGGGTAGGGTA-----GGGG	2687
Db	2429	ATATCAGAGCTTTGAAACAGGTAGTGAATGAATCTCTCCGGGATGGGGTAAACTGGGGTCG	2488



Db	4649	AA	4700
Qy	4675	GCCACACGACTTTTTTAAGATTTTTTATGCCTCTCGATGAGGGAATTTAGTCAATCTATCCT	4734
Db	4709	AA	4768
Qy	4735	CGTCTATTTTGGTGGCTTCTCCGTATTTTAAATTTCTAGTTTGGCACTCCCTTCTCTGAGAG	4794
Db	4769	AA	4828
Qy	4795	CACGCGCTTCGACGAGTAGTTAATACTCTGAGGGCAGGCTTCTGTGAAAGGTTGGCTCG	4854
Db	4829	AA	4888
Qy	4855	GCTCAGTGTGAGATTTTCCCATAAAAAGGGCTCTGCGCCCTGTGTACAGACAGATCGAA	4914
Db	4889	AA	4948
Qy	4915	TCTAGAGTGCACTACGAGATGCCCGGGTTCCGGGGCTCTGATCTCAGGGCATCTTTGC	4974
Db	4949	AA	5008
Qy	4975	CTAGAGATCCTCTACGCCGAGCGATCGTGGCCGGGTACCGAGCTCGAATTCGTAATCAT	5034
Db	5009	AA	5068
Qy	5035	GGTCATAGCTCTTTCTGTGTGAAATTTGTATCCGCTCACAAATCCACACAAATACGAG	5094
Db	5069	AA	5128
Qy	5095	CCGGAAGCATAAAGTGTA-----AG	5115
Db	5129	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAATTTGTCGAAGCTTGGCGAGATTTTCAGGAG	5188
Qy	5116	CCTGGGTGCTTAATGATGAGCTAACTCACATTAATTTGGTTGGCTCACTGCC-----	5171
Db	5189	CTAAGGAAGCTAAAAATGGAGAAAAAATCACCTGGATATACACCGTTTGATATATCCCAAT	5248
Qy	5172	-----GCTTTCCAGTCGGGAAACCTGT	5193
Db	5249	GGCATCGTAAGAACAATTTTGAGGCAATTCAGTCAGTTGCTCAATGTACTATAACAGA	5308
Qy	5194	CGTCCAGCTGCATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTTCGTATTGGGC	5253
Db	5309	CCGTTCACTGCAATTAATGAATCGGCCAAACCGCGGGGAGAGCGGTTTTCGTATTGGGC	5368
Qy	5254	GCTCTTCGGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCGCGAGCGG	5313
Db	5369	GCTCTTCGGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCGCGAGCGG	5428
Qy	5314	TATCAGCTCACTCAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGCAAGAA	5373
Db	5429	TATCAGCTCACTCAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGCAAGAA	5488
Qy	5374	AGAAATGTGACAAAAAGGCCAGCAAAGCCAGGAAACCGTAAAAAGGCCGCTTCTCTCG	5433
Db	5489	AGAAATGTGACAAAAAGGCCAGCAAAGCCAGGAAACCGTAAAAAGGCCGCTTCTCTCG	5548
Qy	5434	GGTTTTTCCATAGGCTCGGCCCGCTGACAGAGATCAAAAAATCGACGCTCAAGTCAGA	5493
Db	5549	GGTTTTTCCATAGGCTCGGCCCGCTGACAGAGATCAAAAAATCGACGCTCAAGTCAGA	5608
Qy	5494	GGTGGCAAAACCGACAGGACTATAAGATACACAGCGCTTTCCCTCGGAGCTCCCTCG	5553
Db	5609	GGTGGCAAAACCGACAGGACTATAAGATACACAGCGCTTTCCCTCGGAGCTCCCTCG	5668
Qy	5554	TGCGCTCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCGG	5613
Db	5669	TGCGCTCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCGG	5728
Qy	5614	GAACGTGGCGCTTTCTCATAGCTCAAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTC	5673
Db	5729	GAACGTGGCGCTTTCTCAATGCTCAAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTC	5788

Qy	5674	GCTCAAGCTGGGTGTGTGCAAGAAACCCCGCTTCAGCCGACCGCTGCGCCTTATCCG	5733
Db	5789	GCTCAAGCTGGGTGTGTGCAAGAAACCCCGCTTCAGCCGACCGCTGCGCCTTATCCG	5848
Qy	5734	GTAACATCGTCTTGTAGTCCAAACCCGGTAAGACAGACTTATCGCCACTGGCAGCAGCA	5793
Db	5849	GTAACATCGTCTTGTAGTCCAAACCCGGTAAGACAGACTTATCGCCACTGGCAGCAGCA	5908
Qy	5794	CTGGTAAACAGGATTAGCAGACGCGAGGTATGTAGGCGGTGCTACAGAGTCTTCTGAAGTGGT	5853
Db	5909	CTGGTAAACAGGATTAGCAGACGCGAGGTATGTAGGCGGTGCTACAGAGTCTTCTGAAGTGGT	5968
Qy	5854	GGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTCGCTCTCTCTGAAGCCAG	5913
Db	5969	GGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTCGCTCTCTCTGAAGCCAG	6028
Qy	5914	TTACCTTTCCGAAAAGAGTTGGTAGCTCTTTGATCCGGCAAAACAAACACCGCTGGTAGCG	5973
Db	6029	TTACCTTTCCGAAAAGAGTTGGTAGCTCTTTGATCCGGCAAAACAAACACCGCTGGTAGCG	6088
Qy	5974	GTGGTTTTTTTGTGTGCAAGCAGCAGATTAACGGCAGAAAAAAGGATCTCAAGAAGATC	6033
Db	6089	GTGGTTTTTTTGTGTGCAAGCAGCAGATTAACGGCAGAAAAAAGGATCTCAAGAAGATC	6148
Qy	6034	CTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAAACTCACTGAAGGGATTT	6093
Db	6149	CTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAAACTCACTGAAGGGATTT	6208
Qy	6094	TGGTCATGAGATTATCAAAAAGGATCTTGACCTAGATCTTTTTAAATTAATAATGAAGTT	6153
Db	6209	TGGTCATGAGATTATCAAAAAGGATCTTGACCTAGATCTTTTTAAATTAATAATGAAGTT	6268
Qy	6154	TTAAATCAATCTAAAGTATATATAGTAAACTTGGTCTGACAGTACCBAATGCTTAATCA	6213
Db	6269	TTAAATCAATCTAAAGTATATATAGTAAACTTGGTCTGACAGTACCBAATGCTTAATCA	6328
Qy	6214	GTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCTGACTCCCCG	6273
Db	6329	GTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCTGACTCCCCG	6388
Qy	6274	TCGTGTAGATAACTACGATACGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATAC	6333
Db	6389	TCGTGTAGATAACTACGATACGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATAC	6448
Qy	6334	CGGAGACCAACGCTCACCGGCTCCAGATTATACGCAATTAACACGACGACGCGAAGG	6393
Db	6449	CGGAGACCAACGCTCACCGGCTCCAGATTATACGCAATTAACACGACGACGCGAAGG	6508
Qy	6394	CCGAGCGCAAGTGGTCTGCAACTTTATTCGCCCTCCATCCAGTCTATTAATTTGTGCC	6453
Db	6509	CCGAGCGCAAGTGGTCTGCAACTTTATTCGCCCTCCATCCAGTCTATTAATTTGTGCC	6568
Qy	6454	GGGAAGCTAGATTAAGTGTTCGCCAGTTAAATGATTTTGGCAACAGTTTGGCAATTCGTA	6513
Db	6569	GGGAAGCTAGATTAAGTGTTCGCCAGTTAAATGATTTTGGCAACAGTTTGGCAATTCGTA	6628
Qy	6514	CAGGC-TCGGTGTGTACCGCTCGTCTGTTGGTATGGCTTCATTCAGCTCCGGTCCCAAC	6572
Db	6629	CAGGCATCGTGTGTACCGCTCGTCTGTTGGTATGGCTTCATTCAGCTCCGGTCCCAAC	6688
Qy	6573	GATCAAGGCGAGTTACATGATCCCGCATGTTGTGCAAAAAGCGGTAGTCTCTTCGGTTC	6632
Db	6689	GATCAAGGCGAGTTACATGATCCCGCATGTTGTGCAAAAAGCGGTAGTCTCTTCGGTTC	6748
Qy	6633	CTCCGATCGTGTGTGAGAAAGTGTGGCCCGAGTGTTCATCTCATGGTTATGGCAGCAC	6692
Db	6749	CTCCGATCGTGTGTGAGAAAGTGTGGCCCGAGTGTTCATCTCATGGTTATGGCAGCAC	6808
Qy	6693	TGCATAATCTCTTACTGTATGCAATCCGTAAGATGCTTTTCTGTGATCTGGTGAATCT	6752
Db	6809	TGCATAATCTCTTACTGTATGCAATCCGTAAGATGCTTTTCTGTGATCTGGTGAATCT	6868

```

6753 CAACCAAGTCATCTGAGAAATAGTGTATGCGGACCGAGTTCCTCTTCCCGCGCTCAA 6812
6869 CAACCAAGTCATCTGAGAAATAGTGTATGCGGACCGAGTTCCTCTTCCCGCGCTCAA 6928
6813 TAGGGATTAATCCGCGGCACATAGCAGAACTTTAAAGTGTCTCATCTGGAACGTT 6872
6929 TAGGGATTAATCCGCGGCACATAGCAGAACTTTAAAGTGTCTCATCTGGAACGTT 6988
6873 CTTCCGGGGGAAACCTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCCA 6932
6989 CTTCCGGGGGAAACCTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCCA 7048
6933 CTCGTGACCCCACTGATCTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAA 6992
7049 CTCGTGACCCCACTGATCTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAA 7108
6993 AAACAGGAAGGCAAAATCGCGCAAAAAGGGAATAAGGGCGACACGCAAAATGTTGAATAC 7052
7109 AAACAGGAAGGCAAAATCGCGCAAAAAGGGAATAAGGGCGACACGCAAAATGTTGAATAC 7168
7053 TCATACCTCTCTCTTTTCAATATATTGAGCAATTTATCAGGGTTATTTCTCATGAGCG 7112
7169 TCATACCTCTCTCTTTTCAATATATTGAGCAATTTATCAGGGTTATTTCTCATGAGCG 7228
7113 GATACATATTGAAATGATTTAGAAAAATAAACAATAGGGTTCCGCGCACATTTCCCC 7172
7229 GATACATATTGAAATGATTTAGAAAAATAAACAATAGGGTTCCGCGCACATTTCCCC 7288
7173 GAAAGTGCACCTGACGCTGAAGAAACCATTTATATCATGACATTAACCTATAAATA 7232
7289 GAAAGTGCACCTGACGCTGAAGAAACCATTTATATCATGACATTAACCTATAAATA 7348
7233 GCGGTATCAGGCGGCTTTCGTC 7256
7349 GCGGTATCAGGCGGCTTTCGTC 7372

RESULT 8
AB086385
LOCUS AB086385 6277 bp DNA circular SYN 14-NOV-2003
DEFINITION Retroviral vector pCX4neo DNA, complete sequence.
ACCESSION AB086385
VERSION AB086385.1 GI:26449191
KEYWORDS Retroviral vector pCX4neo
SOURCE Retroviral vector pCX4neo
ORGANISM other sequences; artificial sequences; vectors.
REFERENCE
1 Akagi, T., Sasai, K. and Hanafusa, H.
Refractory nature of normal human diploid fibroblasts with respect
to oncogene-mediated transformation
Proc. Natl. Acad. Sci. U.S.A. 100 (23), 13567-13572 (2003)
MEDLINE 22975258
PUBMED 14597713
REFERENCE 2 (bases 1 to 6277)
AUTHORS Akagi, T.
DIRECT SUBMISSION
Submitted (10-JUN-2002) Tsuyoshi Akagi, Osaka Bioscience Institute,
Molecular Oncology; Furuedai 6-2-4, Suita, Osaka 565-0874, Japan
(E-mail: takagi@obi.or.jp, Tel: 81-6-6872-4834, Fax: 81-6-6871-7521)
Location/Qualifiers
1. 6277
/organism="Retroviral vector pCX4neo"
/mol_type="other DNA"
/db_xref="taxon:197851"
1. 466
/notes="CMV IE enhancer"
467..762
/notes="R-US region of MuLV 5' LTR"
763..1667
/notes="psi+; extended packaging signal"
1668..2045
/notes="Mo-MuLV pol region containing splice acceptor"

FEATURES
source
enhancer
LTR
misc_signal
misc_feature

```

```

misc_feature 2044..2106
RBS /note="multiple cloning site"
2107..2716
gene /note="IRES sequence from EMCV"
2718..3512
CDS /gene="neo"
2718..3512
/gene="neo"
/codon_start=1
/product="aminoglycoside phosphotransferase"
/protein_id="BAC41724.1"
/db_xref="GI:26449192"
/translation="MIEQDGLHAGSPAAWVERLFGYDWAQQTIGSDAAVFLRLSAQGR
PVLFKTDLGALNELODEAARLSWLTTPVCAAVLDVVTAGRDWLLGLGVPQDL
LSSHLAPAEKYSIMADARLHTLOPATCPDPQAKHRIERARTMEAGLVDDDDLE
EHQGLAPAEFLARLKARPDGDELVTTHGDAACLPNIMVENGFRSFGIDCGRGLGVADRY
QDIALATRDIAELGGEWADRFVLVYIGIAAPDSQRIAFYRLLEDF"
3572..4163
LTR /note="3' LTR of Mo-MuLV"
complement(5216..6076)
gene /gene="bla"
complement(5216..6076)
CDS /gene="bla"
/codon_start=1
/product="beta-lactamase"
/protein_id="BAC41725.1"
/db_xref="GI:26449193"
/translation="NSIOHFRVALIPFFAAFLPVFAHPETLVKVKDAEDQLGARVGY
IELDNSKILESFRPEFRPMWSTFKLLGVLRVDSVDAEQGLRRIRHYSQNDLVE
YSPVTEKLTIDGMTVRELCSAAITMSDNTAANLLTTTGGPKELTAFLHNMGDHVTPL
DRWPELAEFIAPNDRDTPMPVAMATTKLLTGLLTLASQQILDMWEADKVGPL
LRSALPAGWFIADKSGAGERSGIIAALGPDGKPSRIWVIVTTGSAQTMDSNRQIA
ETGASLIKHW"

ORIGIN
Query Match 41.5%; Score 3179.2; DB 12; Length 6277;
Best Local Similarity 73.1%; Pred. No. 0;
Matches 4719; Conservative 0; Mismatches 1048; Indels 691; Gaps 22;

QY 843 TTCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCCAGGACCTGAAATGACCTGTGCCT 902
DB 466 TGCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCCAGGACCTGAAATGACCTGTGCCT 525
QY 903 TATTTGAACCTAACCAATCAGTTTCGCTTCTGCTTCGCTTCGCTTCGCTTCGCTTCGCT 962
DB 526 TATTTGAACCTAACCAATCAGTTTCGCTTCTGCTTCGCTTCGCTTCGCTTCGCTTCGCT 585
QY 963 CTCAATAAAGAGCCCAACCCCTCCTCAGGCGGCCAGTCTCCGATAGACTGCGTCCG 1022
DB 586 CTCAATAAAGAGCCCAACCCCTCCTCAGGCGGCCAGTCTCCGATAGACTGCGTCCG 645
QY 1023 CCGGGTACCGGTATTCGCAATAAAGCCTCTGCTGTTTGGCATCGAATCGTGGACTCGC 1082
DB 646 CCGGGTACCGGTATTCGCAATAAAGCCTCTCTTGCAG-ATTGCATCCGACTTGTGGTCTCGC 704
QY 1083 TGATCCTTGGGAGGGTCTCCTCAGATTGATTGACTGCCC-ACCTCGGGGTCTTTTCATTT 1141
DB 705 TGTTCCTTGGGAGGGTCTCCTCTGAGTGTATGACTACCCGTCAGCGGGGTCTTTTCATTT 764
QY 1142 GGAGGTTTCCACGAGATTTTGGAGACCCCTCGCCAGGAGACCACCGACCCCGCCCGGGAG 1201
DB 765 GGGGGCTCGTCCGGGATCGGAGACCCCTCGCCAGGAGACCACCGA-CCCAACACCGGGAG 823
QY 1202 GTAAGCTGGCCAGCAACTTATCTGTGCTGTCGAGTTGTCTAGTGTCTATGATGATTTT 1261
DB 824 GTAAGCTGGCCAGCAACTTATCTGTGCTGTCGAGTTGTCTAGTGTCTATGATGATTTT 883
QY 1262 ATGCGCTGCGTCCGTTAGTGTAGTAACTAGTCTGTATCTGGCGGACCCCGTGTGGA 1321
DB 884 ATGCGCTGCGTCCGTTAGTGTAGTAACTAGTCTGTATCTGGCGGACCCCGTGTGGA 943
QY 1322 ACTGACGAGTTCCGGAACACCCCGCCGCAACCTCGGAGACGTCCTCGGAGACTTCGGGGC 1381

```

Db 944 ACTGACGAGTTTCGGAACACCCCGCGCAACCTCGGAGACGTCCTCCAGGACCTTCGGGGC 1003  
Qy 1382 GCTTTTGTGCGCGACCTGAGTCTTAAATCCCGATCGTTTAGGACTCTTTGTGTGAC 1441  
Db 1004 GCTTTTGTGCGCGACCTGAGTCTTAAATCCCGATCGTTTGTGAGTCTTTGTGTCAC 1063  
Qy 1442 CCCCTTAGGAGGATATGTTCTGCTGAGGAGACGAGAACCTTAAACAGTTTCCCGC 1501  
Db 1064 CCCCTTAGGAGGATATGTTCTGCTGAGGAGACGAGAACCTTAAACAGTTTCCCGC 1123  
Qy 1502 TCCGCTGCTGAAATTTTGTCTTTCGGTTTGGGACCGAGCGCGCGCTCTTGTCTGCTG 1561  
Db 1124 TCCGCTGCTGAAATTTTGTCTTTCGGTTTGGGACCGAGCGCGCGCTCTTGTCTGCTG 1183  
Qy 1562 CAGCATCGTCTGTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1621  
Db 1184 CAGCATCGTCTGTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1243  
Qy 1622 CCCGGCTAGACTGTTACCACTCCCTTAAGTTTTCAGCTTAGGTCACTGGAAGATGCGA 1681  
Db 1244 CC-----AGACTGTTACCTCCCTTAAAGTTTTCAGCTTAGGTCACTGGAAGATGCGA 1297  
Qy 1682 GCGGATCGCTCAACCACTCGGTAGATGTTCAAGAGAGAGAGTTGGGTACCTTCTGCTC 1741  
Db 1298 GCGGATCGCTCAACCACTCGGTAGATGTTCAAGAGAGAGAGTTGGGTACCTTCTGCTC 1357  
Qy 1742 TGCAGATGGCCAACTTTTAACTGCGGATGCGCGAGACGCGCACTTTAAACCGAGACCT 1801  
Db 1358 TGCAGATGGCCAACTTTTAACTGCGGATGCGCGAGACGCGCACTTTTAAACCGAGACCT 1417  
Qy 1802 CATACCCAGGTTAAGATCAAGGCTTTTTCAGCTGCGCGGATGCGACACCGAGCT 1861  
Db 1418 CATACCCAGGTTAAGATCAAGGCTTTTTCAGCTGCGCGGATGCGACACCGAGCT 1477  
Qy 1862 CCCCTACATCGTACCTGCGAAGCTTGGCTTTTGAACCCCTCCCTGGGTCAAGCCCTT 1921  
Db 1478 CCCCTACATCGTACCTGCGAAGCTTGGCTTTTGAACCCCTCCCTGGGTCAAGCCCTT 1537  
Qy 1922 TGTAACCTTAAGCTCCGCTCTCTTCTCCATCCGCGCGCTCTCTCCCGCTTGAAC 1981  
Db 1538 TGTAACCTTAAGCTCCGCTCTCTTCTCCATCCGCGCGCTCTCTCCCGCTTGAAC 1597  
Qy 1982 TCCTGTTGACACCGGCTCGATCGCTCTTATTCAGCGCTCTCTCTCTCTAGGGC 2041  
Db 1598 TCCTGTTGACACCGGCTCGATCGCTCTTATTCAGCGCTCTCTCTCTCTAGGGC 1657  
Qy 2042 CCCCATATGGCCATATGAGATCTTATATGGGGCACCCCGCGCTTGAACCTTCCCTGA 2101  
Db 1658 CCCCATATGGCCATATGAGATCTTATATGGGGCACCCCGCGCTTGAACCTTCCCTGA 1717  
Qy 2102 CCCTGACATGACAGAGTTACTTAACAGCGCTCTCTTCCAGCTCACTTACAGGCT-TCTA 2160  
Db 1718 CCCTGACTTGACAGAGTTACTTAACAGCGCTCTCTTCCAGCTCACTTACAGGCTCTCTA 1777  
Qy 2161 CTTAGTCCAGCAAGAGTTCTGGAGCTCTGGGGAGCCTTCAAGAACAACTTGGACCG 2220  
Db 1778 CTTAGTCCAGCAAGAGTTCTGGAGCTCTGGGGAGCCTTCAAGAACAACTTGGACCG 1837  
Qy 2221 ACCGGTGTACTCACCTTACCGAGTCCGGGACACAGTGTGGGTCCGCGACACAGAC 2280  
Db 1838 ACCGGTGTACTCACCTTACCGAGTCCGGGACACAGTGTGGGTCCGCGACACAGAC 1897  
Qy 2281 TAAGAACCTAGAACCTCTGCTGGAAGGACCTTACACAGTCTCTGTGACACCCCGACCGC 2340  
Db 1898 TAAGAACCTAGAACCTCTGCTGGAAGGACCTTACACAGTCTCTGTGACACCCCGACCGC 1957  
Qy 2341 CCTCAAAGTAGACGAGCTTGCAGCTTGGATACACCGCGCCCACTGGAAGGTGCGGACCC 2400  
Db 1958 CCTCAAAGTAGACGAGCTTGCAGCTTGGATACACCGCGCCCACTGGAAGGTGCGGACCC 2017  
Qy 2401 CGGGGTGGACCATCTCTAGACTGCC-----ATGGGATGAGCTGTATCATCTCT 2452  
Db 2018 CGGGGTGGACCATCTCTAGACTGCCCGGATCCAGTGTGGGTAGCGGAAATTCGGAC 2077

Qy 2453 TCCTTGTAGCAACAGCTACAGGTGTCTCACTCCGACATCCAGCTGACCCAGAGCCCAAGCA 2512  
Db 2078 CGACCGGTGCGCGCGCGGTAAACAATTGTTTAACTTAAGCTAGCAACGGTTTCCCTCT 2137  
Qy 2513 GCCTGAGCGCGCAGCGTGGTGACAGAGTGACCATCACCTGTAAGCCAGTCAGGATGTGG 2572  
Db 2138 AGCGGATCAATTCGCGCCCCCCCCCTTAACTGCTGCGCGAAGCCCTTGGAAATAGG 2197  
Qy 2573 GTACTTCTGTAGCTTGGTACAGCAGAACGCAAGGTCCAAAGCTGCTGATCTACT 2632  
Db 2198 CCGGTGTGCGTTTGTCTATATGTTTATTTCCACCATATTTGCGCGTCTTTTGGCAATGTGAG 2257  
Qy 2633 GGACATCCACCCCGCACACTGCTGTGCAAGCAGATTTCAGCGGTAGCGGTACCG 2692  
Db 2258 GGCCCGGAAACCTGCGCTCTTCTTTCAGCAGCATTCCTAGGGGTCTTTTCCCTCTCGC 2317  
Qy 2693 ACTTCACCTTCACCATCAGCAGCCCTCCAGCCAGAGGACATCGCCACCTACTTCTGCCAGC 2752  
Db 2318 CAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGAAAGAGTTCCTCTTGGAAAGCTTCTTG 2377  
Qy 2753 AATATAGCCTCTATTCGGTCTG-----TTCCGGCCAAAGGACCAAGGTGGAATCAAAACG 2804  
Db 2378 AAGACAACAACTGCTGTAGCGACCTTTGCAGCAGCGGAACCCCCACCTGCGGACAG 2437  
Qy 2805 AGTGGCTCAGGATCGGGTGGATCCGCTCTGTGGTCTAGGATCGGAGGTCCAACTGGT 2864  
Db 2438 GTGCTCTGCGGCCAAAGACCACTGTATAGATACACCTGCAAAAGCGGCACAAACCCA 2497  
Qy 2865 GGAGAGCGGTGGAGTGTGTGCAACCTGCGCGCTCTGCGCTGCTCTGCTCGGATC 2924  
Db 2498 GTGCCAGTGTGAGTTGGATAGTTGTGGAAGAGTCAAAATGGGTCTCTCTCAAGCGTAT 2557  
Qy 2925 TGGCTTCGATTTACCAACATATTTGGATGAGTTGGGTGAGACAGGACCTCTGAAAGAGTCT 2984  
Db 2558 CAACAAGGGCTGAGGATGCCAGAGGTACCCCATTTGATGGAATCTGATCTGGGGCC 2617  
Qy 2985 TGAAGTGGATGGAAATTCATCCAGATAGCAGTACGATTAACCTATGCGCGTCTCTAAA 3044  
Db 2618 TCGGTGCACATGCTTTACGTGTGTTAGTCAGGTAAATAAAACGCTCTAGGCCCGCAAC 2677  
Qy 3045 GGATAGATTTACATATTCGGAGACACCGCAAGCAACATTTCTCTGCA----- 3096  
Db 2678 CAGCGGACGTGGTCTTCTTTTGAATAACAGATATCCATGATGAACAAAGATGGAT 2737  
Qy 3097 --ATGGACAGCTCAGACCCGAGACACCGGGTCTATTTTGTGCAAGCTTTTACTTCG 3154  
Db 2738 GCAGCAGGTTCTCCGCGCGCTTGGGTGGAGAGCTATTCGGCTATGACTGGGCACACA 2797  
Qy 3155 G--CTTCCCTGGTCTTATTGGGGCCAAAGGACCCCGTCAACCGTCTCCAGTGTAA 3212  
Db 2798 GACAATCGGTCTCTGATGCGCGGTGTTCCGCTGTGACGCGAGGGCGCCCGTCT 2857  
Qy 3213 GCCACACAGACCGCGCGGACCAACCAACCGCGGCCACCATCGCGTGCAGCC 3272  
Db 2858 TTTTGTCAAGACCGACCTGCTCGGTGCGCTGAATGAACCTGACGAGCAGGCGCGCT 2917  
Qy 3273 CTTGCTCCCGCCAGAGCGGCTCGCGCAGCGCGGGCGGCGCAGTGCACAC--GAG 3329  
Db 2918 ATCGTGTGCGCACGAGCGGCTTCTTGGCGAGCTGTGCTCGACGTTGCTACTGAAAGC 2977  
Qy 3330 GGGCTGTGATCTTCCCTGGATCCCAAACTCTGTACTCTGTGATGGAATCTCTTCAT 3389  
Db 2978 GGGAGGACTGTGCTGCTATTGGCGAAGTCCCGGGCAGGATCTCTGTCTACTCTCCT 3037  
Qy 3390 CTATGTTGCTATCTCATCTGCTTGTCTGAGAGTGAAGTTTCAGAGGAGCGGACGCC 3449  
Db 3038 TGCTCTCGCGAGAAAGTATCCATCATGGGTGATGCAATGCGCGGTGCTGATACGCTTGA 3097  
Qy 3450 CCGCGCTACACAGCGGCGAGACCACTTAAACGAGCTCAATCTAGGACCAAGAGA 3509  
Db 3098 TCGGCTACTTGGCCATTCGACCAACCAAGGAAACATCGCATCGAGGAGCAGTACTCG 3157



Qy	3510	GGAGTACGATGTTTTGGACAAAGACAGTGGCCGGGACCTTGAGATGGGGGAAAGCCGAG	3569
Db	3158	GATGGAAGCCGGTCTTCTCGATCAGGATGATCTGGACGAAGAGCATCAGGG--GCTCGCG	3215
Qy	3570	AAGGAAGAACCCCTCAGGAAGCCCTGTACAATGACTCTCAGAAAGATGAAGATGCGGAGGC	3629
Db	3216	CCAGCCGAACCTGTTCGCCAGGCTCAAGGCGCGCATGCCGACGGCAGGATCTCGTCGTG	3275
Qy	3630	CTACAGTTCAGATTGGGATGAAAGCGGACGCGCGAGGGGCAAGG--GGCACGATGGCCCTT	3687
Db	3276	ACCATGGCGATGCTCTTCCGCAATATCATGGTGGAAATGGCCGCTTTTCTGGATTTC	3335
Qy	3688	TACAGGGTCTCAGTACAGCAACCAAGACACCTCAGACGCCCTTCACTCAGAGGCCCTG	3747
Db	3336	ATCGACTGTGGCCGGCTGGGTGTGGCGGACCGGTATCAGGACATAGCGTTGGCTACCCGT	3395
Qy	3748	CCCCCTCGCTAA----CTCGACCGGCGCGGATCCGGATTCGGATTAGTCCAAATTTGTTAAGAC	3803
Db	3396	GATATTCCTGAAGAGCTTTGGCGGCGAATGGGCTGACCGCTCTCTCGTGTCTTACGGTATC	3455
Qy	3804	AGGATATCAGTGGTTCAGGCTCTAGTTTTCGACTCAACAAATATCACCAAGCTGAAGCCCTATA	3863
Db	3456	GCCGCTCCGATTCGACGACGATCGCCCTCTATCGCCTTCTTGACGAGTCTTCTGAGCG	3515
Qy	3864	GAGTACGAGCCATAGATAAAATAAAAGATTTTATTAGTCTCCAGAAAAAGGGGGGAATG	3923
Db	3516	GGACTCTGGGTTTCGATAAAATAAAAGATTTTATTAGTCTCCAGAAAAAGGGGGGAATG	3575
Qy	3924	AAAGACCCCACTGTAGGTTTGGCMAGCTAGCTTAAGTAAAGCCATTTTTCAGAGGCATGG	3983
Db	3576	AAAGACCCCACTGTAGGTTTGGCAAGCTAGCTTAAGTAAAGCCATTTTTCGAAAGGCATGG	3635
Qy	3984	-AAATAATATACTGAGAAATAGAGAAGTCTCAGATCAAGGTTTAGGAAACAGA-GAGACAGCA	4041
Db	3636	AAATAATATACTGAGAAATAGAGAAGTCTCAGATCAAGGTTTAGGAAACAGATGGAACAGCT	3695
Qy	4042	GAATATGGGCAAAACAGATATCTTGTTAAGCAGTTTCCTGCCCC--GCTCAGGGCCAAAGA	4100
Db	3696	GAATATGGGCAAAACAGATATCTTGTTAAGCAGTTTCCTGCCCCCGGCTCAGGGCCAAAGA	3755
Qy	4101	ACAGTTGGAACAGAGATATGGGCCAAACAGGATATCTGTGTTAAGCAGTTTCTGCCCC	4160
Db	3756	ACAGATGAACAGCTGGAATATGGGCCAAACAGGATATCTGTGTTAAGCAGTTTCTGCCCC	3815
Qy	4161	GGCTCAGGGCCAAACAGATGTTCCCCAGATGGTCCCGGCTCCCGGCTCAGGAGTTCTAGAG	4220
Db	3816	GGCTCAGGGCCAAACAGATGTTCCCCAGATGGTCCCGGCTCAGGAGTTCTAGAG	3875
Qy	4221	AACCATCAGATGTTTCCAGGGTCCCCAAAGACCTGAAATGACCTGTGCTTATTGTAA	4280
Db	3876	AACCATCAGATGTTTCCAGGGTCCCCAAAGACCTGAAATGACCTGTGCTTATTGTAA	3935
Qy	4281	CTAACCAATCAGTTCGCTTCTCGCTTCTGTTTCGGCGCTTCTGCTCCCCGAGCTCAATAA	4340
Db	3936	CTAACCAATCAGTTCGCTTCTCGCTTCTGTTTCGGCGCTTCTGCTCCCCGAGCTCAATAA	3995
Qy	4341	AAGAGCCCAACACCCCTCACTCGCGCGCCAGTCCCTCCGATAGACTCGTCCGCCGGGTA	4400
Db	3996	AAGAGCCCAACACCCCTCACTCGCGCGCCAGTCCCTCCGATAGACTCGTCCGCCGGGTA	4055
Qy	4401	CCCGTGTCTCAATAAAACCCCTCTTTCAGTTGCAATCCGACTCGTGGTCTCGCTGTCCCTTG	4460
Db	4056	CCCGTGTCTCAATAAAACCCCTCTTTCAGTTGCAATCCGACTCGTGGTCTCGCTGTCCCTTG	4115
Qy	4461	GGAGGGTCT-CTCTGAGTGATTGACTACCGGTGAGCGGGGTCTTTCAGTTTCTCCCACT	4519
Db	4116	GGAGGGTCTCTCTGAGTGATTGACTACCGGTGAGCGGGGTCTTTCAGTTTCTCCCACT	4155
Qy	4520	ACACAGGTCTCACTAAACATTCCTGATGTGCGCAGGAGCACTCCGTGAGCCGGTTTTTGTT	4579
Db	4156	-----	4155
Qy	4580	TATAATAAATGCAAGAACAGTGTTCCTTTCAGGCCAGACTACATCCTGACTCTCGGCTT	4639

[illegible]





```

2137..2757
/note="IRES sequence from ECMV"
7531..3545
/gene="neo"
2751..3545
/feature="neo"
/function="neomycin resistance"
/codon_start=1
/protein_id="BAA94847.1"
/db_xref="GI:7649703"
/translation="MISQDGHAGSPAAWVERLFGYDWAQQTIGCSDAAPFLRSAGQR
PVLFFVTDLSGALNELQDEARLARLSWLTATGVCFAVLDDVVTFRAGRDMLLGEVPGQDL
LSSHKLPAEKYSIWADARLHTLDDPATGPHQAKHRIERATRMEAGLVDDQDDLDE
EHQGLPAELFAELKARMPGDEGDIIVVTHGDACLPIINVENRFGSGFTDCRGLGVADRY
ODIALATRDIAEELGGEWADRFLVLYGIAAPDSQRIAPFYRLDDEFF"
3605..4197
/note="3' LTR of MuLV"
complement(5187..6047)
/feature="bla"
complement(5187..6047)
/feature="bla"
/function="ampicillin resistance"
/codon_start=1
/product="beta-lactamase"
/protein_id="BAA94848.1"
/db_xref="GI:7649704"
/translation="MSTQHFRVALI PFFAACFLPVFAHPETLVKVKDAEDQLGARVY
I ELDLNSKKILLESFRRPEFPMSTFVKLLCGAVLSRVDAQOQLGRRIHYSQNDLVE
YSPVTERKHLTDGMTVRELCSAAITWSDNTAANLLLTITGGKELTAFHNNGHVGRLL
DRWPELNEALIPNDERTTTPMVAWATTLRKLTGELLTLASRQQLIDMDEADKVAQGL
LRSLALPWGFIADKSGAGERGSGRGIILALGPDGPKPFRIVVITYTGSQATMDERNRQIA
EIGASLIKHW"

```

## ORIGIN

Query Match	41.0%;	Score 3140;	DB 12;	Length 6248;
Best Local Similarity	72.8%;	Pred. No. 0;		
Matches 4720;	Conservative 0;	Mismatches 990;	Indels 776;	Gaps 21;
Qy	843	TTCTAGAGAACCATCAGATGTTTCCAGGGTGCCCCAAGGACCTGAATGACCCCTGTGCCT	902	
Db	466	TGCTAGAGAACCATCAGATGTTTCCAGGGTGCCCCAAGGACCTGAATGACCCCTGTGCCT	525	
Qy	903	TATTTGAATTAACCAATCAGTTTCGCTTCTCGCTTCTGTTCGCGGCCTTCGTCTCCCGGAG	962	
Db	526	TATTTGAATTAACCAATCAGTTTCGCTTCTCGCTTCTGTTCGCGGCCTTCGTCTCCCGGAG	585	
Qy	963	CTCAATAAAGAGCCCAACACCCCTCACTCGGCGCGCCAGTCTCTCCGATAGACTGCGTCG	1022	
Db	586	CTCAATAAAGAGCCCAACACCCCTCACTCGGCGCGCCAGTCTCTCCGATAGACTGCGTCG	645	
Qy	1023	CCGGGTACCCGTATTCGTAATGAAGCCTCTTGCTGTTTCATCCGGAATCGTGGACATCCG	1082	
Db	646	CCGGGTACCCGTGATGCCAATAAACCCCTCTTGCAG--TTGCATCCGACTTGTGGTCTCGC	704	
Qy	1083	TGATCCTTGGAGGGGTCTCTCAGATTGATTGACTGCCC--ACCTCGGGGGTCTTTTCATTT	1141	
Db	705	TGTTCTTGGAGGGTCTCTCTGAGTGATTGACTACCCGTCAGCGGGGGTCTTTTCATTT	764	
Qy	1142	GGAGGTTCCACCGAGATTTGGAGACCCCTGCCCCAGGACCAACGACCCCCCGCGGGAG	1201	
Db	765	GGGGGCTCGTCCGGGATCGGGAGACCCCTGCCCCAGGACCAACGA--CCCAACACCGGGAG	823	
Qy	1202	GTAAGCTGGCCAGCAACTTATCTGTGTCGTGCCGATGTCTAGTGTCATGACTGATTTT	1261	
Db	824	GTAAGCTGGCCAGCAACTTATCTGTGTCGTGCCGATGTCTAGTGTCATGACTGATTTT	883	
Qy	1262	ATGCGCTCGGTCGGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGGAGCCCGTGTGGA	1321	
Db	884	ATGCGCTCGGTCGGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGGAGCCCGTGTGGA	943	
Qy	1322	ACTGACGAGTTCGGAAACACCGGCGCGAAACCTCGGGAGAGCTCCAGGGAGCTTCGGGGGC	1381	
Db	944	ACTGACGAGTTCGGAAACACCGGCGCGAAACCTCGGGAGAGCTCCAGGGAGCTTCGGGGGC	1003	

QY 2461 GCACAGCTACAGGTGTCCTCCAGCATCCAGCTCAGCTGACCCAGAGCCCAAGCAGCCTGAGC 2520  
Db 2084 GGTACCCGGGGCGCGCGGATCTTAAATTAAGCGCGCGGTAACTATTGTTAACTAAC 2143  
QY 2521 -----GCCAGCGTGGGTGACAGAGTACCATCACCTGTAAAGGCACTCAGGATGTGGG 2573  
Db 2144 TAACTAGCTAGCAACGGTTTCCCTCTAGCGGGATCAATTCGGCCCCCCCCCTAACGT 2203  
QY 2574 TACTTCTGTAGCTTGGTACCAGCAGAAAGCCAGGTAAGGCTCAAAGCTGCTGATCTACTG 2633  
Db 2204 TACTGGCCGAAGCGCTTGGAAATAGGCGCGTGTGCGTTGCTATATGTTATTTTCCAC 2263  
QY 2634 GACATCCACCC-----GGCACACTGGTGTGCAAGCAGATTCAGCGGTAGCGGTAGCGGTA 2689  
Db 2264 CATATTGCGCTTTTGGCAATGTGAGGGCCCGGAACCTGGCCCTGCTGCTTTGACGAG 2323  
QY 2690 CCGACTTCACCTCAGCAGCTCCAGCCAGAGGACATCGCCACCTACTACTGCCC 2749  
Db 2324 CATTCCTAGGGTCTTTCCCTCTCGCCAAAGGAATGCMAGGTCTGTGAATGTCGTGA 2383  
QY 2750 AGCAATATAGCTCTATCGGTGTTGCGCCCAAGGACCAAGGTGGAATCAAAAGAGGTG 2809  
Db 2384 GGAAGCAGTCTCTGGAAGCTCTTGAAGACAAACAACTCTGTAGCGACCTTTTGCAG 2443  
QY 2810 GCTCAGGATCGGTGGATCCGGCTCTGTGCTCAGGATCGAGGTCCAACCTGTGGAGA 2869  
Db 2444 GCAGCGGAACCCCCACCTGCGCGACAGGTGCTCTGCGGCCAAAGCCACGTGATAGA 2503  
QY 2870 -----GGGGTGGAGTGTGTGCAACCTGCGCGGTCCCTGCGCTGCTCTCGGATC 2924  
Db 2504 TACACCTGCAAGGGGCAACCCAGTGCACGTTGTGTGATGTGATGTGTGGAAG 2563  
QY 2925 TGGCTTCTGATTACCAATATGGAATGAGTTGGGTGAGACAGGCACCTGGAAGAGTCT 2984  
Db 2564 AGTCAATGGCTCTCTCAAGCGTATTCAACAAGGGCTGAAGATGCCAGAGGTACC 2623  
QY 2985 TGAATGATGGAGAAATATCCAGATAGCAGTACGATTAACATATGCGCGCTCTCTAAA 3044  
Db 2624 CCATTGTATGGATCTGATCTGGGGCTCTCGGTGCAC-ATGCTTTACGTGTGTTTGTGCGA 2682  
QY 3045 GGATAGATTTTACATATATCGCAGACAGCCCAAGACACATGTTCTCTGCAAA----- 3097  
Db 2683 GGTAAATAAAGCTTAGGCCCCCGAACAAGGGGACGTGTTTCTCTTTGAAATAACAG 2742  
QY 3098 TGGACAGCTCAGACCCGAGACACCGGGTCTATTTTGTGCAAGCCTTTACTTTCGGCT 3157  
Db 2743 ATATACCATGATGAACAAGATGGAATGACAGGTTCTCCGGCGCTTGGGTGGAGA 2802  
QY 3158 TCCCTGTTGTTGTTATTTGGGGCCAAAGGACCCCGGTCAAGTCTCCAGTGTCTAAGCCCA 3217  
Db 2803 GGCTATTGGGTATGACTGGGCAACAAGACAAATCGGCTGCTCTGATGCGCGCTGTTC 2862  
QY 3218 CCAGAGCCAGCGCGGACCAACAACCGCGGCCACCATCGCTGCGAGCCCTGT 3277  
Db 2863 GGCTGTACGCGCAGGGGCGCGGTCTTTTGTGCAAGACGACTGTGCGGGTCCCTGA 2922  
QY 3278 CCCTGCGCCCAAGAGCGG-----CTCGCCAGCGCGGGGGGGCGAGTGCACA 3325  
Db 2923 ATGAACCTGCAGACGAGCGCGGCTATCTGTGCTGGCCACAAGCGGCTTCTTGG 2982  
QY 3326 CGAGGGGCTGGAATTCGCCCTGGATCCCAAACTCTGCTACTCTGATGGAATCTCT 3385  
Db 2983 CAGCTGTGCTCGAGTGTCTACTGAAGCGGGAAGGACTGCTGCTATTTGGGCGAAGTGC 3042  
QY 3386 TCATCTATGGTGTCTATCTCTGCTGCTGAGTGAAGTTCAGCAGGAGCGCAG 3445  
Db 3043 CGGGGCGAGGATCTCTCTGATCTCACTCTGCTGCTGCGCGGAGAAAGTATCCATCATGCTG 3102  
QY 3446 AGCC-----CCCGGCTTACACGAGGGCCAGAACAGCTCT 3481  
Db 3103 ATGAATGCGCGGCTGATAGCTGTTATGCTCGGCTACTGCGCCATTGCAACCAAGGGA 3162  
QY 3482 ATACGAGCTCAATCTAGGACGAAGAGAGGAGTACGATGTTTTGGACAAGAGAGCTGGCC 3541

Db 3163 AACATCGATCGAGCGACGTAATCTCGGATGGAAGCCGGTCTTGTGATCAGGATGATC 3222  
QY 3542 GGGACCTTGAGATGGGGGAAAGCCGAGAGGAAGAAACCTTCAGAGAAAGCCCTGTACAATG 3601  
Db 3223 TGGAGGAAGACATCAGG--GCTCGCGCCAGCCGAACTGTTCCGCCAGCTCAAGCGCG 3280  
QY 3602 AACTGCAGAAAGATAAGATGGCGGAGCCTTACGTGAGATTTGGGATGAAAGCGCAGCC 3661  
Db 3281 CATGCCGAGCGGAGGATCTCGTGTGACCATGCGGATGCCCTGCTTCCGGAATATCAT 3340  
QY 3662 GGAGGGGCAAGG--GGCACGATGGCCTTTACCAGGTCTCAGTACAGCCACCAAGGACAC 3719  
Db 3341 GGTGAAATATGGCGCTTTTCTGGAATTCATGCACTGTGGCGGCTGGGTGTGGCGGACCG 3400  
QY 3720 CTAGACGCCCTTCACATGACGCGCTCGCCCTCGCTAA-----CTCGACCGCGCGCGG 3775  
Db 3401 CTATCAGACATAGCTTGGCTACCTGATATATGCTGAGAGCTTGGCGGGAATGGGC 3460  
QY 3776 ATCCGGAATTAGTCCAAATTTGTTAAAGACAGGATATCAGTGTGTCAGGCTCTAGTTTGA 3835  
Db 3461 TGACCGCTTCTCTGCTTTACGGTATCGCGCTCCCGATTCGCGAGGCAATCGCTTCTA 3520  
QY 3836 TCAACATATACACAGCTGAAGCCTATAGAGTACGAGCCATAGATTAATAAAGATTTT 3895  
Db 3521 TCGCTTCTTGAAGCTTCTTGAAGGCTCTGAGCGGACTCTGGGGTTTCGATTAATAAAGATTTT 3580  
QY 3896 ATTTAGTCTCCAGAAAAGGGGGAATGAAGACCCCACTGTAGTTTGGCAAGCTAGC 3955  
Db 3581 ATTTAGTCTCCAGAAAAGGGGGAATGAAGACCCCACTGTAGTTTGGCAAGCTAGC 3640  
QY 3956 TTAAGTAACGCCATTTTGAAGGCTATGGAATAATACATACTGAGATAGAGAGTTTCA 4014  
Db 3641 TTAAGTAACGCCATTTTGAAGGCTATGGAATAATACATACTGAGATAGAGAGTTTCA 3700  
QY 4015 ATCAGGTTAGGAACACAGAG-ACAGCAGATATATGGGCCAAACAGGATATCTGTGTAAG 4073  
Db 3701 ATCAAGTTCAGGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGTAAG 3760  
QY 4074 CAGTTCCTGCCCC--GCTCAGGGCCAAAGACAGTTTGAACAGGAGAAATATGGGCCAAACAG 4132  
Db 3761 CAGTTCCTGCCCCGCTCAGGGCCAAAGACAGATGGAACAGCTGAATATGGGCCAAACAG 3820  
QY 4133 GATATCTGTGTAAGCAGTTCTGCCCCGCTCAGGGCCAAAGACAGATGTCCTCCAGAT 4192  
Db 3821 GATATCTGTGTAAGCAGTTCTGCCCCGCTCAGGGCCAAAGACAGATGTCCTCCAGAT 3880  
QY 4193 GCGTCCCGCTCAGCAGTTCTAGAGAACCATCAGATGTTTCCAGGGTCCCCARAGGA 4252  
Db 3881 GCGTCCAGCCCTCAGCAGTTCTTAGAGAACCATCAGATGTTTCCAGGGTCCCCARAGGA 3940  
QY 4253 CCTGAATGACCTGTGCTTATTTGAACCAATCAATCAGTTCGCTTCTCGCTTCTGTTTC 4312  
Db 3941 CCTGAATGACCTGTGCTTATTTGAACCAATCAATCAGTTCGCTTCTCGCTTCTGTTTC 4000  
QY 4313 GCGGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGCGCGCGCAG 4372  
Db 4001 GCGGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGCGCGCGCAG 4060  
QY 4373 TCCTCCGATAGCTGCTGCTGCGGGTACCGGTTCCTCAATAAACCCTCTTGCAGTTGC 4432  
Db 4061 TCCTCCGATGAGTGAAGTCCCGGGTACCGGTTCCTCAATAAACCCTCTTGCAGTTGC 4120  
QY 4433 ATCCGACTCTGCTGCTGCTTCCCTTGGGAGGGTCT-CTCTGATGATGATGATGATGATGAT 4491  
Db 4121 ATCCGACTCTGCTGCTGCTTCCCTTGGGAGGGTCTCTCTCTGATGATGATGATGATGATGAT 4180  
QY 4492 CAGCGGGTCTTTCAGTTCCTCCACCTACACAGGTCTCACTAACAATTCCTGATGTGCGC 4551  
Db 4181 CAGCGGG----- 4188  
QY 4552 CAGGACTCCTCAGCCCGGTTTTTTTATATAAATAAAGAACAGTGTTCCTTCA 4611



Qy	6771	AATAGTGTATCGCGGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGC	6833
Db	5762	AATAGTGTATCGCGGACCGAGTTGCTCTTGCCCGGCGTCAACACGGGATAATACCGCGC	5821
Qy	6831	CACATAGCAGAACTTTAAAGTGTCTCATCTTGGAAAACGTTCTTCGGGGGGAACCTCT	6890
Db	5822	CACATAGCAGAACTTTAAAGTGTCTCATCTTGGAAAACGTTCTTCGGGGGGAACCTCT	5881
Qy	6891	CAAGGATCTTACCGCTGTTGAGATCCAGTTTCGATGTAAACCCACTCGTGCACCCAACTGAT	6950
Db	5882	CAAGGATCTTACCGCTGTTGAGATCCAGTTTCGATGTAAACCCACTCGTGCACCCAACTGAT	5941
Qy	6951	CTTCAGCATCTTTTACTTTCACACGCGTTCTCGGTGAGCAAAACAGGAAGCAAAATG	7010
Db	5942	CTTCAGCATCTTTTACTTTCACACGCGTTCTCGGTGAGCAAAACAGGAAGCAAAATG	6001
Qy	7011	CCGCAAAAAGGGAATAAGGCGACACGGAAATGTTGAACTACTCATCTCTCTCTTTTC	7070
Db	6002	CCGCAAAAAGGGAATAAGGCGACACGGAAATGTTGAACTACTCATCTCTCTCTTTTC	6061
Qy	7071	AATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAAATGA	7130
Db	6062	AATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAAATGA	6121
Qy	7131	TTTAGAAAAATAAACAAATAGGGTTCCGCGACATTTCCCGAAAAAGTGCCACCTGACG	7190
Db	6122	TTTAGAAAAATAAACAAATAGGGTTCCGCGACATTTCCCGAAAAAGTGCCACCTGACG	6181
Qy	7191	TCTAAGAAACCATTTATTATCATGACATTAACCTATAAAATAGGCGTATCACGAGGCCCT	7250
Db	6182	TCTAAGAAACCATTTATTATCATGACATTAACCTATAAAATAGGCGTATCACGAGGCCCT	6241
Qy	7251	TTTCGTC 7256	
Db	6242	TTTCGTC 6247	
RESULT 10			
CQ879091			
LOCUS	CQ879091	7257 bp	DNA linear PAT 11-OCT-2004
DEFINITION	Sequence 4 from Patent WO2004062554.		
ACCESSION	CQ879091		
VERSION	CQ879091.1 GI:54033468		
KEYWORDS	synthetic construct		
SOURCE	other sequences; artificial sequences.		
ORGANISM	1		
REFERENCE	Martinez-Serrano,A., Liste,I. and Villa,A.		
AUTHORS	Enhancement of neuron generation and survival		
TITLE	Patent: WO 2004062554-A 4 29-JUL-2004;		
JOURNAL	NSGENE A/5 (DK)		
FEATURES	Location/Qualifiers		
source	1..7257		
	/organism="synthetic construct"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32630"		
	/note="Viral vector"		
ORIGIN			
Query Match	40.9%	Score 3130.2;	DB 6; Length 7257;
Best Local Similarity	69.6%;	Pred. No. 0;	
Matches 4939;	Conservative 0;	Mismatches 1763;	Indels 391; Gaps 36;
Qy	550	TTTGAAAGACCCACCTGTAGTTTGGCAAGCTAGCTTAAGTAAACGCAATTTTGCAGGC	609
Db	167	TGTGAAAGACCCACCTGTAGTTTGGCAAGCTAGCTTAAGTAAACGCAATTTTGCAGGC	226
Qy	610	ATGG-AAAAATACATACTGAGATAGAGAAGTTTCAGATCAAGCTTAGGACACA-GAGAC	667
Db	227	ATGGAAAAATACATACTGAGATAGAAAAGTTTCAGATCAAGCTTAGGACACAAGATGGAAC	286
Qy	668	AGCAGAAATATGGGCAACAGGATATCTGTGGTAAGCAGTTCTCTGCCCCC-GCTCAGGGCC	726



Db 1365 GAATGGCCACCTTTAAACGTGCGATGGCCGCGAGACGGCACCTTTAAACGAGACCTCATC 1424  
 QY 1806 ACCAGGTTAAGATCAAGGTCTTTTCACTGCGCGCATGGACACCCAGACAGGTCCGCC 1865  
 Db 1425 ACCAGGTTAAGATCAAGGTCTTTTCACTGCGCGCATGGACACCCAGACAGGTCCGCC 1484  
 QY 1866 TACATCGTGAACCTGGGAAGCTTGGCTTTTGAACCCCTTCCCTGGGTCAAGCCCTTTGTA 1925  
 Db 1485 TACATCGTGAACCTGGGAAGCTTGGCTTTTGAACCCCTTCCCTGGGTCAAGCCCTTTGTA 1544  
 QY 1926 CACCTAAGCTCGGCTCTCTTCTCCATCGGCGCGCTCTCTCCCTTGAACCTCTCT 1985  
 Db 1545 CACCTAAGCTCGGCTCTCTTCTCCATCGGCGCGCTCTCTCCCTTGAACCTCTCT 1604  
 QY 1986 CGTTGACCCGCGCTCGATCTCTCTTTATCCAGCCCTCACTCTCTCTAGCGCGCC 2045  
 Db 1605 CGTTGACCCGCGCTCGATCTCTCTTTATCCAGCCCTCACTCTCTCTAGCGCGCC 1664  
 QY 2046 ATATGGCCATATGAGATCTTATATGGGCAACCCCGCCCTTGTAACTTCCCTGACCT 2105  
 Db 1665 ATATGGCCATATGAGATCTTATATGGGCAACCCCGCCCTTGTAACTTCCCTGACCT 1724  
 QY 2106 GACATGACAGATTACTAACAGCCCTCTCTCCAGCTCTCTTACAGCTCTCTACTTTA 2164  
 Db 1725 GACATGACAGATTACTAACAGCCCTCTCTCCAGCTCTCTTACAGCTCTCTACTTTA 1784  
 QY 2165 GTCCAGCAGAGTCTGGAGACCTCTGGCGGAGCTTACCAAGACCACTGACGACCG 2224  
 Db 1785 GTCCAGCAGAGTCTGGAGACCTCTGGCGGAG - CGTACCAAGAACAACTGACGACCG 1843  
 QY 2225 GTGGTACCTCACCTTACCGAGTGGGAGACACAGTGTGGGTCCGCGACACAGACTAAG 2284  
 Db 1844 GTGGTACCTCACCTTACCGAGTGGGAGACACAGTGTGGGTCCGCGACACAGACTAAG 1903  
 QY 2285 AACCTAGAACCTCTAGACTGGATGGAGCTGTATCATCTCTCTTGTGGAGCAA 2464  
 Db 1904 AACCTAGAACCTCTAGACTGGATGGAGCTGTATCATCTCTCTTGTGGAGCAA 2082  
 QY 2465 CAGCTACAGGTCTCACTCCGACATCCAGTACCCAGAGCCCAAGCAGCTGAGCGCA 2524  
 Db 2083 CGGCCGCGAGTGTCTGGAGCCCATCTCTATTATAAAATGTCTCAGAGCAACCGGAGC 2142  
 QY 2525 GGTGGGTGACAGGTGACC - ATCAGCTGTAGGCCAGTCAAGATGTGGGTACTTCTGTA 2583  
 Db 2143 TGTGGTGGTACTTCTCTCTCAAGCTCTCCAGAAAGGATACAGCTGGAGTCAAGTTTA 2202  
 QY 2584 GCTTGTGTACAGCAGAGCAGGTAAAGGTCCAAAGCTGTGATCTACTTGGACATCCAC 2643  
 Db 2203 GCGATGTCGAGAGACAGGACTGAAGCCCGCAGAGAACTGAACCAAGAGGAGACC 2262  
 QY 2644 CGGCACACTGGTGTGCAAGCAGATTACGCGGTAGCGGTAGCGGTACCGACTTCACTTC 2703  
 Db 2263 CCAGTGCCATCAATGCAACCCATCTCGCACCTGGCGGATAGCCCGCGGTGAATGGAG 2322  
 QY 2704 ACCATCAGCGCTCCAGCCAGAGGACATGCGCACCTACTACTTCCAGCAATATAGCCTC 2763  
 Db 2323 CCATCGGCCACAGCAGAGTTTGGATGCGCGGAGGTACTCCCCATGGCAGCAGTGAAGC 2382  
 QY 2764 TATCGGTGCTTGGCCCAAGGACCAAGGT-----GGAAATCAACGAGGTG 2810  
 Db 2383 AGCGCTGAGAGAGGCTGGCGATGAGTTTGAAGTGGGTACCGGAGAGCATTCAGTGATC 2442  
 QY 2811 CTCAGGATCGGTTGATCCGGCTCTGGTCTCAGGATCGGAG - GTCCAACTGGTGGAGA 2869  
 Db 2443 TAACATCCAGCTTCATATAACCCAGGAGACAGTATATCAGAGCTTTTGAACAGGTAGTGA 2502

QY 2870 GCGGTGAGGTGTTGTGCAACCTGCGCGTCCCTGCGCTGTCTCTGCTCCGATCTGGCT 2929  
 Db 2503 ATGAACTCTTTCGGGATGGGTAAACTGSGGTGCGATTTGTGGCTCTCTCTTTGGGG 2562  
 QY 2930 TCGATTTACACATATTTGGATGAGTGTGGGTGAGACAGGCACTTGAAAGGTCTTGAGT 2989  
 Db 2563 GGGCACTGTGCGTGAAGCGGTAGAACAGGAGATGCAAGTATTTGGTGTGAGTTCGAA 2622  
 QY 2990 GGATTTGGAGAAATTCATCCAGATAGCACTAGATTAACATATGCGCGCTCTCTAAAGGATA 3049  
 Db 2623 G--TTGGATGGGCACTTACCTGATGACCTAGAGCTTGGATCCAGGAGAACGGCG 2680  
 QY 3050 GATTTACATATGCGGAGACCAAGCAACACATTTGCTGCAAAATGGACAGCTGA 3109  
 Db 2681 CTGGGACACTTTTGTGGATCTCTACGGGAACAATACAGCACCCGAGAGCGGAAGGCA 2740  
 QY 3110 GACCCGAGACACCGGGGTCTATTTTTGTGCAAGCTTTACTTTCGGCTTCCCTGGTTG 3169  
 Db 2741 GGAGCGTTTCAACCGCTGTTCTGACGGGATGACTGTGGGTGTGTGTCTATATGTG 2800  
 QY 3170 CTTATTTGGGCAAGGACCCCGGTACCGTCTCCAGTGTCAAGCCACACGAGCGCAG 3229  
 Db 2801 CTCACTTTCAGTTCGGAAGTGACAGATTTGTCAGATATCCATACACTGGCGGCGGAA 2860  
 QY 3230 CGCCCGACCAACCAACCGCGGCCA-----CCATCGGTGCGCAGC 3271  
 Db 2861 TTCCGGATCCACTAGTAAAGCGCGCAGTGTCTGGAATTCGCCCTCTCTCTCCCGCC 2920  
 QY 3272 CCCTGCTCCCTGCGCAGAGCGGTCCGCGCAGCGCGGGGGCGAGTGCAACAGGG 3331  
 Db 2921 CCCTAACGTCTTACCGCGAAGCGCTTGGAAATAGCGCGGTGTGTGTCTATATGTG 2980  
 QY 3332 GGCT-----GGACTTCGCTCGATC 3352  
 Db 2981 ATTTTCCACCATATTTGCCGTCTTTTGGCAATGTGAGGCGCGGAAACCTGGCCCTGTCT 3040  
 QY 3353 CCAAACTCTGTCTACCTGTGTGATGGAATCTCTTCTATGTGTCTATTTCTCACTGCTCT 3412  
 Db 3041 CTTGACGAGCATCTCTAGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTGAA 3100  
 QY 3413 TGTCTTCAGAGTGAAGTTTCAAGG-----AGCGCAGAGCCCCC 3453  
 Db 3101 TGTCTGAGGAAGCAGTTCTCTGAAAGCTTCTTGAAGACAAACAGCTCTGTAGCGAC 3160  
 QY 3454 GGTACACGAGCGGCGAGAACAGCTCTATAAGAGCTCAATCTAGGACGAGAGAGGAG 3513  
 Db 3161 CTTTTCAGGCGAGCGGAAACCCCGACCTGGGACAGGTGCTCTCGCGCAAAAGCCAG 3220  
 QY 3514 TACGATGTTTGGACAGAGACGTGCGCGGACCTGAGATGGGGGGAAGCCG----- 3567  
 Db 3221 TGTATAAGATACCTCTGCAAGCGGCAACCCAGTGCACGTTGTGATGTGATGT 3280  
 QY 3568 -----AGAAGGAAGAACCTCAGGAAGCTGTCAATGAATGCAAGAAAGATAAGAT 3620  
 Db 3281 TGTGGAAGAGTCAAAATGGCTCTCTCAAGCTGTAGTCAACAGGGGCTGAAGATGCCCA 3340  
 QY 3621 GCGCAGGCTTACAG--TGAGATTGGATGAAGG-----GAGCGCGGAGGGGAGGG 3674  
 Db 3341 GAAGGTACCCCATTTATGGGAATCTGATCTGGGGCTCGGTGCAATGCTTTTACATGTG 3400  
 QY 3675 GCACCATGCGCTTTTACAGGGTCTCAGTACAGC - CACCAAGGACACCTACGACGCGCTTC 3733  
 Db 3401 TTTAGTGTGAGGTTAAAGCTCTAGGCGCCCGGAAACCGGGGAGGTGTTTCTTTG 3460  
 QY 3734 ACATGCGAGCCCTGCGCCCTCGCTAACTGACGCGCGCGGATCCGGATTAGTCAAT 3793  
 Db 3461 AAAACACGATGATTAAGCTTGCACAGGAATTCATGTTGAGCAAGCAGATCTTGAAGAAC 3520  
 QY 3794 TGTAAAGACAGGATATCAGTGG--TCCAGGCTCTAGTTTGTGACTCAACAATATCAGCAG 3851  
 Db 3521 ACCGCGCTGCGAGGATCATGAGCTTCAAGGTGAACCTGGAGGCGGTGTGAAACACAC 3580



```
Db 5741 TTGAGTCCACCCGGTAAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGA 5800
QY 5806 TTAGCAGACGAGGTATGTAGGCGGTCTACAGAGTTCTTGAAGTGGTGGCCCTAACTACG 5865
Db 5801 TTAGCAGACGAGGTATGTAGGCGGTCTACAGAGTTCTTGAAGTGGTGGCCCTAACTACG 5860
QY 5866 GCTACACTAAGAGGACAGATTTTGGTATCTGGGCTCTGCTGAAGCCAGTACCTTCGGAA 5925
Db 5861 GCTACACTAAGAGGACAGATTTTGGTATCTGGGCTCTGCTGAAGCCAGTACCTTCGGAA 5920
QY 5926 AAGAGTTGGTAGCTCTTATCGCGCAAAACAAACCAACCGCTGGTAGCGGTGTTTTTTTG 5985
Db 5921 AAGAGTTGGTAGCTCTTATCGCGCAAAACAAACCAACCGCTGGTAGCGGTGTTTTTTTG 5980
QY 5986 TTTGCAAGCAGCAGATTTACGCGCAGAAAGAGATCTCAAGAGATCCTTTGATCTTTT 6045
Db 5981 TTTGCAAGCAGCAGATTTACGCGCAGAAAGAGATCTCAAGAGATCCTTTGATCTTTT 6040
QY 6046 CTACGGGGTCTGACGCTCAGTGGAAACGAAAACTCAGTTTAAAGGATTTTGGTCAAGAT 6105
Db 6041 CTACGGGGTCTGACGCTCAGTGGAAACGAAAACTCAGTTTAAAGGATTTTGGTCAAGAT 6100
QY 6106 TATCAAAAAGGATCTTCACTAGATCCTTTTAAATTAATAATGAAGTTTAAATCAATCT 6165
Db 6101 TATCAAAAAGGATCTTCACTAGATCCTTTTAAATTAATAATGAAGTTTAAATCAATCT 6160
QY 6166 AAGATATATAGTAAACCTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTA 6225
Db 6161 AAGATATATAGTAAACCTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTA 6220
QY 6226 TCTCAGCGATCTGTCTATTTCTTCCATCCATAGTTGCTTGAATCTCCCGCTCGTGTAGATAA 6285
Db 6221 TCTCAGCGATCTGTCTATTTCTTCCATCCATAGTTGCTTGAATCTCCCGCTCGTGTAGATAA 6280
QY 6286 CTACAGATACGGAGGGCTTACCATCTGGCCCGCAGTCTGCAATGATACCGGAGACCCAC 6345
Db 6281 CTACAGATACGGAGGGCTTACCATCTGGCCCGCAGTCTGCAATGATACCGGAGACCCAC 6340
QY 6346 GCTCACCGGCTCCAGATTTATCAGCAATAAACAACGACCGAGCGGAGCGGCAGAA 6405
Db 6341 GCTCACCGGCTCCAGATTTATCAGCAATAAACAACGACCGAGCGGAGCGGCAGAA 6400
QY 6406 GTGGTCTCTGCAACTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAGCTAGAG 6465
Db 6401 GTGGTCTCTGCAACTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAGCTAGAG 6460
QY 6466 TAAGTAGTTCCGCAAGTAAATAGTTTGGCAACGTTGTCGCAATGCTACAGGC-TCGTGG 6524
Db 6461 TAAGTAGTTCCGCAAGTAAATAGTTTGGCAACGTTGTCGCAATGCTACAGGCATCGTGG 6520
QY 6525 TGTACGCTCGTGGTTGTATGGCTTTCATTCAGCTCCGCTCCCAACGATCAAGCGAG 6584
Db 6521 TGTACGCTCGTGGTTGTATGGCTTTCATTCAGCTCCGCTCCCAACGATCAAGCGAG 6580
QY 6585 TTACATGATCCCCCATGTTGTGCAAAAAGCGGTAGCTCCTTGGTCTCCGATCGTTG 6644
Db 6581 TTACATGATCCCCCATGTTGTGCAAAAAGCGGTAGCTCCTTGGTCTCCGATCGTTG 6640
QY 6645 TCAGAAAGTAAAGTTGGCCGAGTGTATCACTCATGTTATGGCAGCACTGCATAATTCCTC 6704
Db 6641 TCAGAAAGTAAAGTTGGCCGAGTGTATCACTCATGTTATGGCAGCACTGCATAATTCCTC 6700
QY 6705 TTACTGTCAATGCCATCCGTAAGATGCTTTTCTGTGACTGGTAGTACTCAACCAAGTCAAT 6764
Db 6701 TTACTGTCAATGCCATCCGTAAGATGCTTTTCTGTGACTGGTAGTACTCAACCAAGTCAAT 6760
QY 6765 TCTGAGAAATAGTGTATGCGCGCACCGAGTTGCTTCTTGGCCGGCTCAATACGGGATAATA 6824
Db 6761 TCTGAGAAATAGTGTATGCGCGCACCGAGTTGCTTCTTGGCCGGCTCAACCGGATATA 6820
QY 6825 CCGCGCCACATACGACAACTTTAAAAAGTGTCTCATCATTTGAAAAACGTTTCTCGGGCGAA 6884
Db 6821 CCGCGCCACATACGACAACTTTAAAAAGTGTCTCATCATTTGAAAAACGTTTCTCGGGCGAA 6880
```

```
QY 6885 AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGACCCA 6944
Db 6881 AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGACCCA 6940
QY 6945 ACTGATCTTCAGCATCTTTTACTTTTACACAGCGTTTCTGGGTGAGCAAAAACAGGAAGC 7004
Db 6941 ACTGATCTTCAGCATCTTTTACTTTTACACAGCGTTTCTGGGTGAGCAAAAACAGGAAGC 7000
QY 7005 AAAATGCCGCAAAAAGGGAATAAGGCGACACGGAATGTTGAATACTCATACTCTTCC 7064
Db 7001 AAAATGCCGCAAAAAGGGAATAAGGCGACACGGAATGTTGAATACTCATACTCTTCC 7060
QY 7065 TTTTTCATATTTATTTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTG 7124
Db 7061 TTTTTCATATTTATTTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTG 7120
QY 7125 AATGATTTAGAAAAATAAACAATAGGGTTCCGCGACATTTCCCGAAAAAGTGCCAC 7184
Db 7121 AATGATTTAGAAAAATAAACAATAGGGTTCCGCGACATTTCCCGAAAAAGTGCCAC 7180
QY 7185 CTGAGCTTAAGAAACCATTTATTCATGACATTAACCTATAAAATAGGCGTATCACGA 7244
Db 7181 CTGAGCTTAAGAAACCATTTATTCATGACATTAACCTATAAAATAGGCGTATCACGA 7240
QY 7245 GGCCCTTTCTGCT 7257
Db 7241 GGCCCTTTCTGCT 7253

RESULT 11
AB086386 6076 bp DNA circular SYN 14-NOV-2003
LOCUS Retroviral vector pCX4pur DNA, complete sequence.
DEFINITION AB086386
ACCESSION AB086386
VERSION GI:26449194
KEYWORDS
SOURCE Retroviral vector pCX4pur
ORGANISM Retroviral vector pCX4pur
REFERENCE 1
AUTHORS Akagi,T., Sasai,K. and Hanafusa,H.
TITLE Refractory nature of normal human diploid fibroblasts with respect to oncogene-mediated transformation
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (23), 13567-13572 (2003)
MEDLINE 22975258
PUBMED 14597713
REFERENCE 2 (bases 1 to 6076)
AUTHORS Akagi,T.
TITLE Direct Submission
JOURNAL Submitted (10-JUN-2002) Tsuyoshi Akagi, Osaka Bioscience Institute, Molecular Oncology; Furuedai 6-2-4, Suita, Osaka 565-0874, Japan (E-mail: takagi@obi.or.jp; Tel:81-6-6872-4834, Fax:81-6-6871-7521) Location/Qualifiers
FEATURES
source
1..6076
/organism="Retroviral vector pCX4pur"
/mol_type="other DNA"
/db_xref="taxon:197852"
enhancer
1..466
/note="CMV IE enhancer"
LTR
467..762
/note="R-U5 region of MuLV 5' LTR"
misc_signal
763..1667
/note="psi+; extended packaging signal"
misc_feature
1668..2045
/note="Mo-MuLV pol region containing splice acceptor"
misc_feature
2044..2106
/note="multiple cloning site"
RBS
2107..2716
/note="IRES sequence from EMCV"
gene
2718..3317
/feature="pur"
CDS
2718..3317
```

LTR  
 gene  
 CDS

```

/gene="pur"
/codon_start=1
/product="puromycin-N-acetyl transferase"
/protein_id="BAC41726.1"
/db_xref="GI:26449195"
/translation="MTBYKPTVRLATRDVPRVAVRTLAAAFADYPATRHVTDPDRHIE
RVTELOELFLTRVLGDLGKVVADGAAVA VMTTPESVEAGVPAEIGPRMAELSGSR
LAAQQOEGLLAPHRPEKPAWPLATVGVSPHQKGLGSVVLPGVZAAERAGVPAFL
ETSAPNLPPFYRGLFVTVADVECPKORATWCMRKPGA"
3371..3962
/note="3' LTR of Mo-MuLV"
/complement(5015..5875)
/gene="bla"
/complement(5015..5875)
/gene="bla"
/codon_start=1
/product="beta-lactamase"
/protein_id="BAC41727.1"
/db_xref="GI:26449196"
/translation="MSIQHFVALIPFFAFLPFAHPETLVKVKDAEDQLGARVGY
IELDNGSKILESFRPRLSPFPMSTFKVLLCGAVLSRDVAGQEQGLRRIHYSONDLVE
YSPVTEKHLDTGMRVRELCSAAITMSDNTAANLLLTIGGPKELTALFHNMGDHTVRL
DRWPELNEAIPNDRDITMEVAMATTLRLKLLTGELLTLASROQLIDWMEADKVGPL
LRSALPAGWFTADKSGAGERSGI IIALGPDGKPSRI VVIYITGTSQATWDERNRQIA
EIGASLIKHW"
  
```

## ORIGIN

Query Match 40.7%; Score 3113; DB 12; Length 6076;  
 Best Local Similarity 73.3%; Pred. No. 0;  
 Matches 4708; Conservative 0; Mismatches 895; Indels 818; Gaps 24;

Qy	843	TTCTAGAGAACCATCAGATGTTTCAGGGTGCCCAAGACCTGAATGACCCCTGTGCCT	902
Db	466	TGCTAGAGAACCATCAGATGTTTCAGGGTGCCCAAGACCTGAATGACCCCTGTGCCT	525
Qy	903	TATTGAACTAACCAATCAGTTCCTCTCGCTTCCTGTTTCGGCGCTTCTGCTCCCGAG	962
Db	526	TATTGAACTAACCAATCAGTTCCTCTCGCTTCCTGTTTCGGCGCTTCTGCTCCCGAG	585
Qy	963	CTCAATAAAGAGCCCAACCCCTCACTCGGCGCGCAGTCCCTCCGATAGACTGCGTCG	1022
Db	586	CTCAATAAAGAGCCCAACCCCTCACTCGGCGCGCAGTCCCTCCGATAGACTGAGTCG	645
Qy	1023	CCCGGGTACCGGTATCCCAATAAAGCCTTTGCTGTTTGTGATCCGAATCGTGACTCGC	1082
Db	646	CCCGGGTACCGGTATCCCAATAAAGCCTTTGCTGTTTGTGATCCGAATCGTGACTCGC	704
Qy	1083	TGATCCTTGGAGAGGTCTCTCAGATTGATTGACTGGCC-ACCTCGGGGGTCTTTCATTT	1141
Db	705	TGTTCTTGGAGAGGTCTCTCTGAGTGATTGACTACCCGTCAGCGGGGGTCTTTCATTT	764
Qy	1142	GGAGGTTCCACCGAGATTGGAGACCCCTGCGCGCGCAGTCCCTCCGATAGACTGCGTCG	1201
Db	765	GGGGGCTGTCGCGGATCGGGAGACCCCTGCGCGCGCAGTCCCTCCGATAGACTGAGTCG	823
Qy	1202	GTAAGCTGGCCAGCAACTATCTGTGTCTGTCGGAITTTCTAGTGTCTATGACTGATTTT	1261
Db	824	GTAAGCTGGCCAGCAACTATCTGTGTCTGTCGGAITTTCTAGTGTCTATGACTGATTTT	883
Qy	1262	ATGGCCCTGGTGGTACTAGTTAGCTAACTAGCTCTGTATCTGTATCTGGCGGACCCGTGGTGA	1321
Db	884	ATGGCCCTGGTGGTACTAGTTAGCTAACTAGCTCTGTATCTGTATCTGGCGGACCCGTGGTGA	943
Qy	1322	ACTGACGAGTTCGGAAACACCGCGCGCAACCTCGGGAGAGCTCCACGGGACTTCGGGGGC	1381
Db	944	ACTGACGAGTTCGGAAACACCGCGCGCAACCTCGGGAGAGCTCCACGGGACTTCGGGGGC	1003
Qy	1382	CGTTTTTGTGCCCGCACTGAGTCTTAAATCCCGATCGTTTAGGACTCTTTTGTGTGACCC	1441
Db	1004	CGTTTTTGTGCCCGCACTGAGTCCAAATCCCGATCGTTTGGACTCTTTTGTGTGACCC	1063
Qy	1442	CCCTTAGAGAGGGATATGTGGTTCCTGGTAGGAGACGAGAACCTTAAACAGTTTCCCGCC	1501

Db	1064	CCCCTAATAGAGGGATATGTGGTCTCTGGTAGGAGACGAGAACCTTAAACAGTTCCCGCC	1123
Qy	1502	TCCGCTGGAATTTTGTCTTTCGGTTTGGGACCGGAGCCGCGCGCTCTTGTCTGCTG	1561
Db	1124	TCCGCTGGAATTTTGTCTTTCGGTTTGGGACCGGAGCCGCGCGCTCTTGTCTGCTG	1183
Qy	1562	CAGCATGTTCTGTTGTCTCTCTGCTGACTGTCTCTGTTTGTCTGTTTGTCTGAAATATGGG	1621
Db	1184	CAGCATGTTCTGTTGTCTCTCTGCTGACTGTCTCTGTTTGTCTGTTTGTCTGAAATATGGG	1243
Qy	1622	CCCGGGCTAGACTGTTTACCACTCCCTTAAAGTTTTCACCTTAGGTCACTGGAAGATGTGCA	1681
Db	1244	CC-----AGACTGTTTACCACTCCCTTAAAGTTTTCACCTTAGGTCACTGGAAGATGTGCA	1297
Qy	1682	GCGGATCGCTCACAAACAGTCGGTAGATGTCAAGAAGAGACGTTGGGTACTCTTCTGCTC	1741
Db	1298	GCGGATCGCTCACAAACAGTCGGTAGATGTCAAGAAGAGACGTTGGGTACTCTTCTGCTC	1357
Qy	1742	TGCAGATGGCCAAACCTTTTAAAGTCGGATGGCCCGGAGACGGACCTTTTAAACCGAGACT	1801
Db	1358	TGCAGATGGCCAAACCTTTTAAAGTCGGATGGCCCGGAGACGGACCTTTTAAACCGAGACT	1417
Qy	1802	CATCACCCAGGTTAAGATCAAGGTCTTTTTCACCTGGCCCGCATGCGACACCCAGACCCAGT	1861
Db	1418	CATCACCCAGGTTAAGATCAAGGTCTTTTTCACCTGGCCCGCATGCGACACCCAGACCCAGT	1477
Qy	1862	CCCCTACATCGTGACCTGGGAAGCCTTGGCTTTTGTGACCCCGCTCCCTGGGTCAAGCCCTT	1921
Db	1478	CCCCTACATCGTGACCTGGGAAGCCTTGGCTTTTGTGACCCCGCTCCCTGGGTCAAGCCCTT	1537
Qy	1922	TGTACACCTTAAGCCTCGGCTCTCTTCTCTCGATTCGCGCCCGCTCTCTCCCGCTTGAACC	1981
Db	1538	TGTACACCTTAAGCCTCGGCTCTCTTCTCTCGATTCGCGCCCGCTCTCTCCCGCTTGAACC	1597
Qy	1982	TCCTCGTTTCGACCCCGCTCGATCCCTCTTATTCAGCCCTCACTCTCTCTCTAGGGCGC	2041
Db	1598	TCCTCGTTTCGACCCCGCTCGATCCCTCTTATTCAGCCCTCACTCTCTCTCTAGGGCGC	1657
Qy	2042	CCCCTATGGCCATATGAGATCTTATATGGGGACCCCGCCCTTGTAAACTTCCCTGA	2101
Db	1658	CCCCTATGGCCATATGAGATCTTATATGGGGACCCCGCCCTTGTAAACTTCCCTGA	1717
Qy	2102	CCCTGACATGACAGAGTTACTAACAGCCCTCTCTCCAAAGCTCACTTACAGGCT-TCTA	2160
Db	1718	CCCTGACATGACAGAGTTACTAACAGCCCTCTCTCCAAAGCTCACTTACAGGCTCTCTA	1777
Qy	2161	CTTAGTCCAGCAGGAAGTCTGGAGACCTCTGGCGGAGCCTTACCAAGAACAACTTGAACCG	2220
Db	1778	CTTAGTCCAGCAGGAAGTCTGGAGACCTCTGGCGGAGCCTTACCAAGAACAACTTGAACCG	1837
Qy	2221	ACCGTGGTACCTCACCCCTTACCGAGTCGGCGACACAGTGTGGGTCCGCGCACACAGAC	2280
Db	1838	ACCGTGGTACCTCACCCCTTACCGAGTCGGCGACACAGTGTGGGTCCGCGCACACAGAC	1897
Qy	2281	TAAGAACCTTAGAACCTCGCTGGAAAGGACCTTACACAGTCTCTGCTGACCCACCCACCGC	2340
Db	1898	TAAGAACCTTAGAACCTCGCTGGAAAGGACCTTACACAGTCTCTGCTGACCCACCCACCGC	1957
Qy	2341	CCTCAAGTAGAGCGGATCCGAGCTTGGATACACCGCGCCCAACGTTGAAGGTCGGGACCC	2400
Db	1958	CCTCAAGTAGAGCGGATCCGAGCTTGGATACACCGCGCCCAACGTTGAAGGTCGGGACCC	2017
Qy	2401	CGGGGGTGGACCACTCTCTAGACTGCCATGGAGTGGATCATCTCTCTTCTTGGTA	2460
Db	2018	CGGGGGTGGACCACTCTCTAGACTGCCGATCCAGTGTGGTGTAGCGGAATTCGGG--	2075
Qy	2461	GCAACACTACAGGTGTCCACTCCGACATCCAGCTGACCCAGAGCCCAAGACGCTTGAGC	2520
Db	2076	---ACGACGGGTGGCGCCGGGTAACTTGTAACTTAAGTAGTAGCAACGGTTTC	2132
Qy	2521	GCAGCGTGGGTGACAGGTGACCATCACTGTAAAGCCAGTCAAGGATGTGGGTACTTCT	2580
Db	2133	CCTCTAGCGGGATCAATTCGCGCCCGCCCGCTTAACTTACT-----	2174

QY 2581 GTAGCTTGTACAGCAGAACGAGTAAGGCTCCAAAGCTGCTGATCTACTGACATCC 2640  
 Db 2175 ---GGCGAAGCGCTTGAATAAGCGCGTGTGGTTGTCTATATGTTATTTCCACC 2231  
 QY 2641 ACCCGGACACATGCTGTGTCAGAGAGANTCAGCGGTAGCGGTACCGGATACCACTTACC 2700  
 Db 2232 ATATGCGCTCTTTTG------GCAATGTGAGGGCCCGGAACCTGGCCCTGCTCTCTTG 2285  
 QY 2701 TTCACCATCAGCAGCTCAGCCAGAGGACATGCGCACCTACTACTGCGCAGCAATATAGC 2760  
 Db 2286 ACAGCATTTCTTAGGGGTCTTTCCCTCTCGCAAGGAATCAAGGCTGTGTGAATGTC 2345  
 QY 2761 CTCATCGCTGTTCGGCCAAAGGACCAAGGTGGAATCAAAACAGGCTGCTCAGATCG 2820  
 Db 2346 GTGA--GGAAGCAGTTCCTCTGGAAGCTTCTTGAAGACAAACAGCTGTAGCACCC 2403  
 QY 2821 GGTGATCCGGTCTGTGGTCTCAGGATCGAGGTCCAACTGTGTGAGAGCGGTGAGGT 2880  
 Db 2404 TTTCAGGCGAGGAAACCCCACTGGCGACAGGTGCTCTGGGCCAAAGCCACGTTG 2463  
 QY 2881 GTTGTCAACCTGGCGGTCCTTCGGCTGTCTGCTCGCATCTGGCTTCGATTTCACTACC 2940  
 Db 2464 TATAAGATACCTCTGCAAAAGGCGGCACAAACCCAGTGCACGTTGTGAGTTGGATGTTG 2523  
 QY 2941 ACATATTGATGAGTTGGGTGAGACAGGCACCTGGAAAGGTCCTTGAGTGGATTGGAGAA 3000  
 Db 2524 TGAAGAAGTCAATGCTCTCTCTCAGCGTATTCMAAGGGGTGAAGATGCCAGA 2583  
 QY 3001 ATTTCATCCAGATGACGATACGATTAACTATGCGCGTCTCTAAAGGATGATTTCATA 3060  
 Db 2584 AGGTACCCCATTTGATGGGATCTGATCTGGGCTCGGTGCATGCTTTTACATGTGTTT 2643  
 QY 3061 TCGGAGACAGCCCAAGAACACATTTGCTCTGCAATGGAAGCAGCTGAGACCCGAGAC 3120  
 Db 2644 AGTCGAGGTTAAAAACGTCCTA-----GGCCCCCGCAACACGCGG 2683  
 QY 3121 ACCGGGCTCTATTTTGTGCAAGCTTTACTTTCGGCTTCCCTGCTTCTTATTTGGGCG 3180  
 Db 2684 GACGTGTTTCTTTTAAAAAACACGATTAATCAATGACCGAGTACAGCCACCG- 2739  
 QY 3181 CAAGGACCCCGTCAACCGTCTCAGTGTAGAGCCCAACACGACGCGCAGCGCGGACCA 3240  
 Db 2740 -----TGCGCTCGCCACCCCGCAGCGTCCCGCGCGG 2775  
 QY 3241 CCAACACCGGCGCCACATCGGTCGAGCCCTGTCTGCTGCGCCCGCAGAGCGGTGCG 3300  
 Db 2776 TAGGCACCTCGCGCGCGTTCGCGGACTACCCCGCCACGCGCCACACCGTCCGACCGG 2835  
 QY 3301 CCAGCGCGGGGCGCAGTGACACAGAGGGGCTGAGCTTCGCCCTGGATCCCAAACTC 3360  
 Db 2836 ACGCCACATCGAGCG----GGTACCGAGCTGCAAGAACTCTTCTACGCGCGTGGG 2891  
 QY 3361 TGCTACTGCTGATGGAATCTCTTCTATCTATGTGTCAATCTCACTGCTTGTTCCTG 3420  
 Db 2892 CTCGACATCGCAAGGTGTGGGTGCGGACGAGCGCGCGGTGCGGTCTGGACCAAG 2951  
 QY 3421 AGAGTGAATTACAGAGAGCGCAGAGCCCGCGGTACAGAGGGGCCAGAACAGCTC 3480  
 Db 2952 CCGGAGAGGCTCGAAGCGGGCGGTGTTGCGCGAGATCGGCGCGCATGCGCGAGTTG 3011  
 QY 3481 TATAACGAGCTCAATCTAGACGAGAGAGGATGAGTGTGTTTGGACAGAGAGCTGGC 3540  
 Db 3012 AGCGGTTCCCGGTCGCCCGCAGCAACAGATGGAAGGCTCTCTGGCGCGCAGCGGCC 3071  
 QY 3541 CGGGACCTCTGAGATGGGGGGAAGCCGAGAGGAAGAACCTCAGGAAGGCTCTGACAT 3600  
 Db 3072 AAGGAGCCCGTGGTT-----CCTGGCCACCGTGGCGTCTC 3109  
 QY 3601 GAATCTCAGAAAGATGAAGATGGCGAGGCGCTCAGTGAGATTGGGATGAAGCGAGCGC 3660  
 Db 3110 GCCCGACCAACAGGCAAGGGTCTGGCGCAGCGCGCTCGTGCTCCCCCGGAGTGGAGCGCGC 3169

QY 3661 CGAGGGGCAAGGGGACAGATGCGCTTTTACAGGGTCTCAGTACAGCCCAAGACACACC 3720  
 Db 3170 CGAGCGCGCGGG-----TGCCCGCTTCTCTGGAGAC 3203  
 QY 3721 TAGCAGCGCTTTCATGTCAGGCGCTGCCCTCGCTAACTCGACGCGCGCGGATCCG 3780  
 Db 3204 TCCGCGCGCGCAACCTCCCTTCTACGAGCGGCTCGGCTTCACTGCGCGCGGATCCG 3263  
 QY 3781 GATTAGTCCAAATTTGTAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTGTATCAAC 3840  
 Db 3264 GAGTCCCGAAGGACCGCGGACCTGGTGCATGACCCGAAAGCCCGTGCCTGACGCTC 3323  
 QY 3841 AATATCACCAGCTGAAGCCTATAGAGTACGAGCCATAGATAAAATAAAGATTTTATTA 3900  
 Db 3324 GACA-----CGATAAAATAAAGATTTTATTA 3351  
 QY 3901 GTCTCCAGAAAAGGGGGAATGAAGACCCCACTGTAGTTTGGCAAGCTAGCTTAAG 3960  
 Db 3352 GTCTCCAGAAAAGGGGGAATGAAGACCCCACTGTAGTTTGGCAAGCTAGCTTAAG 3411  
 QY 3961 TAAAGCCATTTGCAAGGCATGG-AAAATACATACACTGAGAAATAGAGAAGTTTCAGTCAA 4019  
 Db 3412 TAAGCCCATTTTGAAGGCATGGAATAATACATACTGAGATAGAGAAGTTTCAGTCAA 3471  
 QY 4020 GGTAGGAACAGA-GAGACAGCAGATATGGGCCAAACAGGATATCTGTGTAAGCAGTT 4078  
 Db 3472 GGTAGGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGTAAGCAGTT 3531  
 QY 4079 CCGTGGCC- GGTAGGGCCAAAGACAGTTGGAACAGGAGATATGGGCCAAACAGGATAT 4137  
 Db 3532 CCGTGGCCGCTCAGGGCCAAAGACAGATGGAACAGCTGAATATGGGCCAAACAGGATAT 3591  
 QY 4138 CTGTGTAAGCAGTTCTGCGCGCTCAGGGCCAAAGACAGATGCTCCAGAGTCCGCT 4197  
 Db 3592 CTGTGTAAGCAGTTCTGCGCGCTCAGGGCCAAAGACAGATGCTCCAGAGTCCGCT 3651  
 QY 4198 CCGCGCTCAGCAGTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCCAGAGCTCA 4257  
 Db 3652 CCGCGCTCAGCAGTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCCAGAGCTCA 3711  
 QY 4258 AATGACCTGTGCTTTTGAACCTAACCAATCAGTTCGCTCTGCTTCTGCTGCGCG 4317  
 Db 3712 AATGACCTGTGCTTTTGAACCTAACCAATCAGTTCGCTCTGCTTCTGCTGCGCG 3771  
 QY 4318 CTTCGCTCCCGGCTCAATAAAGAGCCCAACCCCTCACTCGGCGCGCAGTCTC 4377  
 Db 3772 CTTCGCTCCCGGCTCAATAAAGAGCCCAACCCCTCACTCGGCGCGCAGTCTC 3831  
 QY 4378 CGATAGACTGGCTCGCGCGGTACCGGTGTTCTCAATAAACCTCTTGCAGTTGCATCG 4437  
 Db 3832 CGATTTGACTGAGTCCCGCGGTACCGGTGTTCTCAATAAACCTCTTGCAGTTGCATCG 3891  
 QY 4438 ACTCGTGTCTGCTGCTTCTTGGAGGGTCT-CTCTGAGTGATTTGACTACCGCTCAGCG 4496  
 Db 3892 ACTCGTGTCTGCTGCTTCTTGGAGGGTCTCTCTCTGAGTGATTTGACTACCGCTCAGCG 3951  
 QY 4497 GGGTCTTTCAGTTTCTCCCACTACAGGTTCTCACTAACATTCCTGATGTCGCGCAGCG 4556  
 Db 3952 GGG----- 3954  
 QY 4557 ACTCGCTCAGCGCGGTTTGTATTATAATAAATGCAAGAAACAGTGTCTCCCTTCAAGCCA 4616  
 Db 3955 ----- 3954  
 QY 4617 GACTACATCTGACTCTCGGCTTTATAAAGAAATGTTGAAGGGCTCTGTGGACTATCTGC 4676  
 Db 3955 ----- 3954  
 QY 4677 CACACGACTTTTAAAGATTTTATGCTCTCTGATGAGGGATTTAGTCAATCTATCTCG 4736  
 Db 3955 -----GTCTTTCATGCGAGCTGATATCAAAATTAATTTGG 3990  
 QY 4737 TCTATTTGCTGGCTTCTCGGTATTTTAAATTTTCTAGTTTGCATCTCCCTTCTCTGAGAGCA 4796





QY 6956 GCATCTTTTACCTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGCGCAAAATGCCGA 7015  
DB 5775 GCATCTTTTACCTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGCGCAAAATGCCGA 5834  
QY 7016 AAAAGGGAATAGGCGGACACGGAATGTTGAATCTCATACTCTTCTTTTCAATAT 7075  
DB 5835 AAAAGGGAATAGGCGGACACGGAATGTTGAATCTCATACTCTTCTTTTCAATAT 5894  
QY 7076 TATTGAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTGAAATGATTTAG 7135  
DB 5895 TATTGAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTGAAATGATTTAG 5954  
QY 7136 AAAAATAAACAATAGGGGTTCCGGGCACATTTCCCGCAAAAGTGCCACTGACGCTCTAA 7195  
DB 5955 AAAAATAAACAATAGGGGTTCCGGGCACATTTCCCGCAAAAGTGCCACTGACGCTCTAA 6014  
QY 7196 GAAACCATTTATCATGACATTAACCTATAAAATAGCGGTATCACGAGGCCCTTTGCT 7255  
DB 6015 GAAACCATTTATCATGACATTAACCTATAAAATAGCGGTATCACGAGGCCCTTTGCT 6074  
QY 7256 C 7256  
DB 6075 C 6075

RESULT 12  
AR028670  
LOCUS AR028670 5364 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 4 from patent US 5858744.  
ACCESSION AR028670  
VERSION AR028670.1 GI:5940643  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 5364)  
AUTHORS Baum, C., Stocking-Harbers, C. and Ostertag, W.  
TITLE Retroviral vector hybrids and the use thereof for gene transfer  
JOURNAL Patent: US 5858744-A 4 12-JAN-1999;  
FEATURES  
Location/Qualifiers  
1..5364  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 40.6%; Score 3111; DB 6; Length 5364;  
Best Local Similarity 92.0%; Pred. No. 0;  
Matches 3414; Conservative 0; Mismatches 210; Indels 87; Gaps 9;

QY 3868 ACAGCCATAGATAAATAAAGATTTTATTAGTCTCCAGAAAAGGGGGGAATGAAG 3927  
DB 1655 ACAGCCATAGATAAATAAAGATTTTATTAGTCTCCAGAAAAGGGGGGAATGAAG 1714  
QY 3928 ACCCCACCTGTAGTTTGGCAAGC-----TAGCTTAAGTAAGCCCAT 3969  
DB 1715 ACCCCACCTGTAGTTTGGCAAGCTAGAGTCGCTTAGCCCTGATAGCCGAGTAACCCAT 1774  
QY 3970 TTTTGAAGGCATGG-AAAATACATACTGAGAATAGAGAAGTTTCAGATCAAGGTTAGGAA 4028  
DB 1775 TTTTGAAGGCATGGAAAATAATCCAAACCAAGAAATAGGGAAGTTTCAGATCAAGGGGGTA 1834  
QY 4029 CAGAGAGACAGCAGAAATATGGGCAAAACAGGATATCTGTGTAAGCAGTTCTTGGCCCGC 4088  
DB 1835 CATGAAAATAGCTTAACGTTTGGGCAAAACAGGATATCTGCGGTGAGCAGTTTTCGGCCCGG 1894  
QY 4089 TCAGGCGCAAGAACAGTGTGGAACAGGAGAAATAGGCCCAAAAC-----AGGATATCTG 4140  
DB 1895 CCGGGGCAAGAACAGATGTCACCCAGTTCACCGCCCGCCCGGCGCCCAAGAACAGATG 1954  
QY 4141 TGGTAAGCATGTTCTTCCCGCGGCTCAGGGCCCAAGAAACAGATGTTCCCGAGATCGCGTCCC 4200  
DB 1955 GTACCCGCGATTTTCGGGCCCGCGCGCCCAAGAACAGATGTTCCCGAGATATGGCCCA 2014

QY 4201 GGCCTCAGCAGTCTTAGAGAACCATCAGATGTTTTCCAGGGTGCCCCAAGGACCTGAAT 4260  
DB 2015 ACCCTCAGCAGTCTTAAAGACCCATCAGATGTTTCCAGGCTCCCCAAGGACCTGAAT 2074  
QY 4261 GACCTGTGCTTATTTGAACATAACCAATCAGTTCGCTTCTCGCTTCTGTTCGCGGCTT 4320  
DB 2075 GACCTGTGCTTATTTGAATTAACCAATCAGCCTGCTTCTCGCTTCTGTTCGCGGCTT 2134  
QY 4321 CTGCTCCCGGAGCTCAATTAAGAGAGCCACACCCCTCAGTCGGCGCGCCAGTCTCCGA 4380  
DB 2135 CTGCTCCCGGAGCTCTATAAAGAGCTCACACCCCTCAGTCGGCGCGCAGTCTCCGA 2194  
QY 4381 TAGACTGTGCTGCGCCGGGTACCCGTGTTCTCAATAAACCCTCTTGCAAGTTGATCCGACT 4440  
DB 2195 TTGACTGAGTCCCGCGGTACCCGTGTTCTCAATAAACCCTCTTGCAAGTTGATCCGACT 2254  
QY 4441 COTGTGCTGCTGCTTCTTGGGAGGGTCT-CTCTGAGTGATTTGACTTACCCGTGAGC-GGG 4498  
DB 2255 COTGTGCTGCTGCTTCTTGGGAGGGTCTCTCTCTGAGTGATTTGACTTACCCGTGAGCGGG 2314  
QY 4499 GTCTTTTCAAGTTTCTCCAGCTACACAGGTCTCACTAAACATTTCTCTGATGTCGCCGAGGAC 4558  
DB 2315 GTCTTTTCAAGTTTCTCCAGCTACACAGGTCTCACTAAACATTTCTCTGATGTCGCCGAGGAC 2374  
QY 4559 TCCGTGAGCCCGGTTTTTGTATATAAATGAAGAACAGTGTTCCTTCAAGCCAGA 4618  
DB 2375 TCCGTGAGCCCGGTTTTTGTATATAAATGAAGAACAGTGTTCCTTCAAGCCAGA 2434  
QY 4619 CTACATCTGACTCTCGGCTTTTATAAAGAAATGTTGAAGGGCTCTGTGGACTATCTGCCA 4678  
DB 2435 CTACATCTGACTCTCGGCTTTTATAAAGAAATGTTGAAGGGCTCTGTGGACTATCTGCCA 2494  
QY 4679 CAGGACTTTTAAAGATTTTATGCTCTCTGAGTAGGGAATTTAGTCAATCTATCTCTGCTC 4738  
DB 2495 CAGGAC-TTTTAAAGATTTTATGCTCTCTGAGTAGGGAATTTAGTCAATCTATCTCTGCTC 2553  
QY 4739 TATTTTGTGCTGCTCTCGTATTTTATAAATTTCTAGTTTGCATCTCTCTCTGAGAGCAG 4798  
DB 2554 TATTTTGTGCTGCTCTCGTATTTTATAAATTTCTAGTTTGCATCTCTCTCTGAGAGCAG 2613  
QY 4799 GCGATTGACAGTAGTTAATCTCTGAGGCGCAGGCTTCTGTGAAAGGTTGCTCGGCTC 4858  
DB 2614 GCGATTGACAGTAGTTAATCTCTGAGGCGCAGGCTTCTGTGAAAGGTTGCTCGGCTC 2673  
QY 4859 AGTGTGAGATTTTGCATATAAAGAGGGTCTCTGCTCTGTGTACACAGACATCGGAATCTA 4918  
DB 2674 AGTGTGAGATTTTGCATATAAAGAGGGTCTCTGCTCTGTGTACACAGACATCGGAATCTA 2733  
QY 4919 GAGTGCATACTCAGAGTCCCGCGGTTCCGGGCTCTGATCTCAGGGCATCTTTTGCCTAG 4978  
DB 2734 GAGTGCATACTCAGAGTCCCGCGGTTCCGGGCTCTGATCTCAGGGCATCTTTTGCCTAG 2793  
QY 4979 AGATCTCTTACCGCGGACGATCTGTGGCGGGTACCGAGCTCGAATTCGTAATCATGCTC 5038  
DB 2794 AGATCTCTTACCGCGGACGATCTGTGGCGGGCATCACCGCGCCACAGGTGCGGTTGCTG 2853  
QY 5039 ATAGCTGTTTC-----CTGTGTGAATTTGTTATCCGCTCACAATTTCCACACA 5085  
DB 2854 GCGCTTATATCCCGACATCACCGATGGGGAAGATCGGCTCGCCACTTCGGGCTCATGA 2913  
QY 5086 ACATACAGAGCCGAAGCATATAAGTCTGAGGCTGCGGCTCTAATGAGT----- 5134  
DB 2914 GCGCTTGTTCGGGCTGGGTATGTTGGCAGGCCCGCGGGGGAGCTGTTGGGGGCCA 2973  
QY 5135 -----GAGCTAACTCACAATTAATTCGTTGCGC 5162  
DB 2974 TCTCTTGCATGCACCATTTCTTGGCGCGGCTGCTCAACGCGCTCAACCTACTACTGG 3033  
QY 5163 TCATGCGCGGCTTTCAGTCCGGGAAACCTGTCTGTCCAGCTGCTGATTAATGAATCGGCCAA 5222  
DB 3034 GCTGCTTCTTAATGAGGAGTGCATAAGGGAGAGCGTCTCTGATTAATGAATCGGCCAA 3093



Db 5254 GGGTGTGGGGCTGGCTTAATATATGCGGCATCAGACGAGATGTTACTGAGATGCACCAT 5313  
Qy 7442 ATCCGCTGTGAATACCGCACAGATCGTAAGCAGAAATACCGCATCAGG 7492  
Db 5314 ATCCGCTGTGAATACCGCACAGATCGTAAGCAGAAATACCGCATCAGG 5364

RESULT 13  
U00220  
LOCUS 5506 bp RNA linear VRL 27-DEC-1993  
DEFINITION Human immunodeficiency virus type 1 Nef protein and neomycin  
phosphotransferase genes, complete cds.  
ACCESSION U00220  
VERSION U00220.1 GI:392790  
KEYWORDS  
SOURCE Human immunodeficiency virus 1 (HIV-1)  
ORGANISM Human immunodeficiency virus 1  
VIRUSES; Retroviral viruses; Retroviridae; Lentivirus; Primate  
lentivirus group.  
REFERENCE 1 (bases 1 to 6506)  
Anderson, S., Shugars, D.C., Swanstrom, R. and Garcia, J.V.  
Nef from primary isolates of human immunodeficiency virus type 1  
suppresses surface CD4 expression in human and mouse T cells  
J. Virol. 67 (8), 4923-4931 (1993)  
93323236  
PUBMED 8331733  
REFERENCE 2 (bases 1 to 6506)  
Shugars, D.C.  
Direct Submission  
Submitted (11-AUG-1993) Diane C. Shugars, UNC-Lineberger Cancer  
Research Ctr, University of North Carolina at Chapel Hill, Manning  
Drive, Chapel Hill, NC 27599-7295 USA  
LOCATION/Qualifiers  
1. .6506  
/organism="Human immunodeficiency virus 1"  
/proviral  
/mol\_type="unassigned RNA"  
/db\_xref="taxon:11676"  
/clone="pLconsnefsN"  
1655..2275  
/codon\_start=1  
/products="HIV-1 consensus Nef protein"  
/protein\_id="AAA03700.1"  
/db\_xref="GI:392792"  
/translation="MGKWSKESVSCWPAVRMRRAEPAEGVGNVRDLKHKAIT  
SNRTATNACWLEAQEEVEGFPVRPQVPLRPMTYKAADVLSHLKRGLEGLIY  
SKRQDILDWVHTQGYFPDQWNTYPPGIRPLTFGWCFLVFPVEPEKVEANEGE  
NNCLLHPMSQHGMDDEPEVLVWKFDSKLAFHFMARELHPYYKDC"  
2698..3492  
CDS  
/codon\_start=1  
/product="neomycin phosphotransferase"  
/protein\_id="AAA03699.1"  
/db\_xref="GI:392791"  
/translation="MIEQDLHAGSPAAWVERLFGVDWAQQTIGCSDAVFRLSAQGR  
PVLVFKVTLDSGLNLEQAEARLSLWATVPCAAVLVDVTEAGRDWLLGLBVGQDL  
LSHLAPAEKVSIMADAMERLTLDPATCFPDQAKHLERARTMEAGLVDDDLDE  
EQGLAPAEFLKARMPDGEDLVVTHGDACLINWVNGRFSFIDCRLGVADRY  
QDIALATRDIAELGEGWADRLVLYGAAPSDQSRIAFYRLLDLDEF"

ORIGIN  
Query Match 40.6%; Score 3104.6; DB 14; Length 6506;  
Best Local Similarity 70.8%; Pred. No. 0;  
Matches 4793; Conservative 0; Mismatches 1474; Indels 504; Gaps 30;  
Qy 549 TTTTGAAGACCCACCTAGTTTGGCAAGCTTAAGTAAAGCCCATTTTGAAGG 608  
Db 174 TTTTGAAGACCCACCCCTAGG--TGGCAAGCTAGCTTAAGTAAAGCCCATTTTGAAGG 231  
Qy 609 CATGG--AAATACATACTGAGNATAGAGAGTTTCAGATCAAGTTAGGAACAGAGAC 667  
Db 232 CATGGAATAATACATACTGAGNATAGAGAGTTTCAGATCAAGTTAGGAACAGAGAC 291  
Qy 668 AGCAGAAATAGGCCCAACAGGATATCTGTGGTAAGCAGTTCCTGCCCC--GCTCAGGGCC 726

Db 292 AGCTGAATA---CCAAACAGGATATCTGTGTAAGCGGTCTCTGCCCCGGCTCAGGGCC 347  
Qy 727 AAGAACAGTTTGAACAGGAGAT-ATGGGCCAAACAGGATATCTGTGTAAGCAGTTCT 785  
Db 348 AAGAACAGATGACAGACGCTGAGTGTATGGGCCAAACAGGATATCTGTGTAAGCAGTTCT 407  
Qy 786 GCCCGGGCTCAGGGCCAAAGAACAGATGTTCCACAGATGCGGTCCCGCCCTCAGCAGTTTC 845  
Db 408 GCCCGGGCTCAGGGCCAAAGAACAGATGTTCCACAGATGCGGTCCAGCCCTCAGCAGTTTC 467  
Qy 846 TAGAAGAACCATCAGATGTTTCCAGGGTGCCTCCCAAGGACCTG-AAATGACCCCTGTGCTTTA 904  
Db 468 TAGTGAATCATCAGATGTTTCCAGGGTGCCTCCCAAGGACCTGAAATGACCTGTACTTTA 527  
Qy 905 TTTGAACATCAACATCAGTTTCGCTTCTGTGTTGCGCGCTTCTGCTCCCCGAGCT 964  
Db 528 TTTGAACATCAACATCAGTTTCGCTTCTGTGTTGCGCGCTTCTGCTCCCCGAGCT 587  
Qy 965 CAATAAAGAGAGCCCAACCCCTCACTCGGGCGCCAGTCCTCCGATAGACTGCGTCGCC 1024  
Db 588 CAATAAAGAGAGCCCAACCCCTCACTCGGGCGCCAGTCCTCCGATAGACTGCGTCGCC 647  
Qy 1025 CGGGTACCCGCTATTCCTCAATAAAGCTCTTGTGTTGATCCGAATCGTGGACTGCTG 1084  
Db 648 CGGGTACCCGCTATTCCTCAATAAAGCTCTTGTGTTGATCCGAATCGTGGACTGCTG 707  
Qy 1085 ATCTCTGGAGAGGTCTCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1144  
Db 708 TTTCTTGGAGAGGTCTCTCTCAGATGATGATGATGATGATGATGATGATGATGATGATG 767  
Qy 1145 GGTTCACCGCAGATTTGGAGACCCCTGCGCCAGGAGCCACCGACCCCGCGCGGAGGTA 1204  
Db 768 GGTCTGTCGGGATTTGGAGACCCCTGCGCCAGGAGCCACCGA-CCACACCGCGGAGGTA 826  
Qy 1205 AGCTGGCCAGCAACTTATCTGTGCTGTGCTCGATGTTCTAGTGTCTATGATGATGATGATG 1264  
Db 827 AGCTGGCCAGCAACTTATCTGTGCTGTGCTCGATGTTCTAGTGTCTATGATGATGATGATG 886  
Qy 1265 CGCTGCGTGGTACTAGTTAGCTAACTAGCTCTGTATCTGCGGACCCGCTGGTGAAC 1324  
Db 887 CGCTGCGTGGTACTAGTTAGCTAACTAGCTCTGTATCTGCGGACCCGCTGGTGAAC 946  
Qy 1325 GACGAGTTTCGGAACACCGCGCGCAACCTCGGAGAGCTCCAGGAGACTTCGCGGGCCGT 1384  
Db 947 GACGAGTTTCGGAACACCGCGCGCAACCTCGGAGAGCTCCAGGAGACTTCGCGGGCCGT 1006  
Qy 1385 TTTTGTGGCCCGACCTGAGTCTCTAAATCCCGATCGTTTAGGACTCTTTTGGTGCACCCCC 1444  
Db 1007 TTTTGTGGCCCGACCTGAGGAGGAGTCTGATGTTGGAATCCGACCCCGTCA----- 1057  
Qy 1445 CTTAGAGGAGGATATGTTGTTCTGTTAGGAGACAGAGAACCTAAACAGTTCCCGCTCC 1504  
Db 1058 -----GGATATGTTGTTCTGTTAGGAGACAGAGAACCTAAACAGTTCCCGCTCC 1107  
Qy 1505 GTCTGAATTTTGTGCTTTCGTTTGGAGCCGAGAGCCGCGCGCTTGT---CTGCTG 1561  
Db 1108 GTCTGAATTTTGTGCTTTCGTTTGGAAACGAGAGCCGCGCTTGTCTGCTGAGCGCTG 1167  
Qy 1562 CAGCATCGTTCTGTTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTT 1621  
Db 1168 CAGCATCGTTCTGTTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTT 1227  
Qy 1622 CCGGGGTAGACTGTTTACACTCCCTTAAGTTTGAACCTTAGTGTACTGGAAGAGATGCA 1681  
Db 1228 CC-----AGACTGTTTACCACTCCCTTAAGTTTGAACCTTAGTGTACTGGAAGAGATGCA 1281  
Qy 1682 GGGGATCGCTCAACACAGTCTGATGTTCAAGAGAGAGAGCTTGGTTTACCTTCTGCTC 1741  
Db 1282 GGGGATCGCTCAACACAGTCTGATGTTCAAGAGAGAGAGCTTGGTTTACCTTCTGCTC 1341  
Qy 1742 TGCAGAAATGGCCCAACTTTTAACTCGGATGGCGGAGACGACCTTTTAAACCGAGACCT 1801

Db 1342 TCGAGATGGCCAACTTTTAAGTCGGATGGCCGGAGACGGCACCTTTAAACCGAGCCT 1401  
Qy 1802 CATACCCAGGTTAAGATCAAGGCTTTTCACTGGCCCGCATGGACACCCAGAGGT 1861  
Db 1402 CATACCCAGGTTAAGATCAAGGCTTTTCACTGGCCCGCATGGACACCCAGAGGT 1461  
Qy 1862 CCCCTACATCTGACCTGGGAAGCTTGGCTTTTGAACCCCTCCCTGGGTCAAGCCCTT 1921  
Db 1462 CCCCTACATCTGACCTGGGAAGCTTGGCTTTTGAACCCCTCCCTGGGTCAAGCCCTT 1521  
Qy 1922 TGTACACCTTAAGCTTCGGCTCTCTCTCTCCATCCGCCCCGCTCTCTCCCTTTGAACC 1981  
Db 1522 TGTACACCTTAAGCTTCGGCTCTCTCTCTCCATCCGCCCCGCTCTCTCCCTTTGAACC 1581  
Qy 1982 TCCTCGTTGACACCCGCTCGATCTCCCTTTATCCAGCCCTCACTCTCTCTAGGCGC 2041  
Db 1582 TCCTCGTTGACACCCGCTCGATCTCCCTTTATCCAGCCCTCACTCTCTCTAGGCGC- 1640  
Qy 2042 CCCCATATGGCCATATGAGATCTTATATGGGGCACCCCGCCCTTGTAAACTTCCCTGA 2101  
Db 1641 -CCGAATTCGGCCACATGGGTGGCAAGTGGTCAAAACGTAGTGTAGTGGATGGCCTGCT 1699  
Qy 2102 CCCTGACATGACAAAGATTTACTAACAGCCCTCTCTCTCCAGCTCACTTTACAGGCTTCTAC 2161  
Db 1700 GTAAGAGAAATGAGCGAGCTGAGCCACGACGAGGGAGTGGAGCAG----- 1751  
Qy 2162 TTAGTCCAGCAGAGTCTGGAGACCTCTGGCGGAGCTTACCAAGAACAACTGGACCGA 2221  
Db 1752 -----TATCTCGAGACCTGGAAACATGAGGCAATTTACAAGTAGCAATACAGACTA 1805  
Qy 2222 CCGGTGTACTCTACCTTTACCGAGTCCGGGACACAGTGGGGTCCGCGACACAGACT 2281  
Db 1806 CCAATGTCTGTGTGCTCGCTGGCTAGAAACACAAAGAGGAGAAAGTGGGTTTCCAGTCA 1865  
Qy 2282 AAGAACCTAGAACTCGCTGGAAGGACCTTTACACAGTCTCTGCTGACACCCACCCGCC 2341  
Db 1866 GACCTCAGTAGCTTTAAGACCAATGACTTACAAGGACGAGTAGACTTTAGCCACTTTT 1925  
Qy 2342 CTCAAAGTAGACGCATCG-----CAGCTTGATACAGCCCGCCACGTGAAGGTGCGCA 2397  
Db 1926 TAAAGAAAGAGGGGAGCTGGAAGGGTTAATTTACTCCCAAAAGACAGATATCTTG 1985  
Qy 2398 CCCCGGGGTGGACATCTCTAGACTGCCATGGATGGAGCTGTATCATCTCTTCTTG 2457  
Db 1986 ATCTGTGGGTCTACCAACACACAGGCTACTTCCCTGATTTGGCAGAACTTACACACGAGC 2045  
Qy 2458 GTACCAACAGCTACAGTGTCCACTCCGACATCAGCTGACCCAGAGCCCAAGCAGCCTG 2517  
Db 2046 CAGGAATCAGATATCCAATGACCTTTGGGTGGTCTTCAAACTAGTACCAAGTTGAGCCAG 2105  
Qy 2518 AGCCACGCGTGGGTGACAGAGTGACCATCACTGTAAAGGCCAGTCAAGGATGTGGGTACT 2577  
Db 2106 AAGAGGTAGAGAGGCCAATGAAGGAGAGAACAACTGCTTGCTACACCTATGAGCCAGC 2165  
Qy 2578 TCTGTAGTGGTACCAAGAGAGAGGTAAGGCTCAAAGCTGTGATCTACTGAGCA 2637  
Db 2166 ATGGGATGGATGACCCAGAGAAAGATTTAGTGTGGAGTTTGACAGCAAACTAGCAT 2225  
Qy 2638 TCCACCCGGCACATGGTGTGCCAAGCAGATTACGCGGTAGCGGTAGCGGTACCGACTTC 2697  
Db 2226 TTCAATCAGATGGCCGAGAGCTGATCCGGAGTACTACAAGACTGCTGAGAAATTCGTTA 2285  
Qy 2698 ACCTTCACTACAGCAGCCTCCAGCCAG-AGGACATCGCCACCTACTACTGCCAGCAATA 2756  
Db 2286 ACTCAGGATCCGCTGTGGATGTGTGCTAGTTAGGGTGTGGAAAGTCCCGAGGCTCCC 2345  
Qy 2757 TAGCCTCTATGGGTGTTGGCCAAAGGACCAAGGTGGAAATCAAAACAGGTGGCTCAGG 2816  
Db 2346 CAGCAGGACAGAGTATGCAAGAGCATGCATCTCAATTTAGTCAGCAACCAAGGTGTGGAAAGT 2405  
Qy 2817 ATCGGGTGGATCCGGCTCTGGTGGCTCAGGATCCGAGGTCCAATGTTGGAGAGCGGTGG 2876  
Db 2406 CCCCAGGCTCCCCAGCGGACAGAGTATGCAAGCATGTCATCTCAATTTAGTCTCCAGAA 2465

Qy 2877 AGGTGTTGTGCAACCTGCGCGGTCCCTGCGCTGTCTGCTCCGATCTGCTTGGATTT 2936  
Db 2466 TAGTCCGCGCCCTAACTCCGCCCCTAACTCCGCCCCTAACTCCGCCCCTAACTTCTC 2525  
Qy 2937 CACCACATATTTGGATGAGTTGGGTGACACAGGACCTGGAAAAGGTCTTGTAGTGGATTGG 2996  
Db 2526 CGCCCCATGGCTGACTAATTTTTTTTATTTATGAGAGGCGCGGCCCTCGGCCCTCG 2585  
Qy 2997 AGAAATTCATCCAGATAGCAGTACGATTAACATATGCGCCGCTCTCTAAAGGATAGATTT-- 3054  
Db 2586 AGCTATTTCCAGAGTAGTGTAGGAGGCTTTTGTGAGGCTTAGGCTTTTGTCAAAGAGCTTG 2645  
Qy 3055 -ACAATATCCGAGACAACCCCAAGAACACATTTGTTCTGCAATGAGACACCTGAGACC 3113  
Db 2646 GGCTGCAGGTGAGGCGGATCTGATCAAGACAGAGATGAGATCGTTTCGATGATGA 2705  
Qy 3114 CGAAGACACCGGGGTCTATTTTGTGCAAGCTTTACTTGGCTTCCCTGGTTGCTTGA 3173  
Db 2706 ACAAGATGGATTGACGACAGGTTCTCCGCGCGCTTGGGTGGAGAGGCTATTTCCGCTATGA 2765  
Qy 3174 TTGGGGCCAAAGGAGACCCCGGTACCGTCTCCAGTGTAAAGCCACACGACCGCCAGCGCC 3233  
Db 2766 CTGGGCAACAGACAAATCGCTCTGATGCCCGCGTGTTCGGCTGTGAGCGAGGG 2825  
Qy 3234 GCGACCAACCAACACCGCGCGCCCACTATCGGTTCGACGCCCTGTCTCCCTGCGCCCGAGAGGC 3293  
Db 2826 GCGCCCGGTTCTTTTGTCAAGACCGACCTGTCTCGGTGCCCTGAATGAATGCAGGACGA 2885  
Qy 3294 GG-----CTCGGCGAGGGCGGGGGCGAGTGACACAGAGGGGGCTGGACTT 3341  
Db 2886 GGCAGCGCGCTATCTGTGGCTGGCCACGACGCGGCTTCTTGGCAGAGTGTGTCTGACGT 2945  
Qy 3342 CGCCTCGGATCCAAACTCTGTCTACCTGTGGATGGAATCCTCTTCTATCTATGTGTCTAT 3401  
Db 2946 TGTCACTGAAGCGGGAAGGACTGGCTGCTATTGGGGAAGTGCCTGGGGCAGGATCTCTCT 3005  
Qy 3402 TCTCACTGCTTGTCTCTGAGAGTGAAAGTTTTCAGCAGGAGCGCAGAGCC----- 3449  
Db 3006 GTCATCTCACTTGTCTCTCGCGAGAAAGTATCCATCATGCTGATGCAATGCGCGGCT 3065  
Qy 3450 -----CCCGCGGTACAGAGGGGCGAGAACAGCTCTATTAACGAGCTCAATCT 3497  
Db 3066 GCATACGCTTGTATCCGCTTACCTGCCATTTCCAGACCAAGCGAAACATCGCATCGAGCG 3125  
Qy 3498 AGACAGAGAGAGAGTACGATGTTTTTGGACAAGACGTGGCGGGACCTTGAGATGGG 3557  
Db 3126 AGCACGTACTCGGATGGAAGCCGCTTGTGATGAGGATGATCTGGACGAAGAGCATCA 3185  
Qy 3558 GGGAAAGCCGAGAAAGAAACCTTCAGGAAGGCTGTACAAATGAATGCAACTGCAGAAAGATAA 3617  
Db 3186 GGG--GCTCGCGCCAGCCGAATGTTCCGCCAGGCTCAAGGCGCATGCCCGAGCGAG 3243  
Qy 3618 GATGGCGGAGGCTTACAGTGAATTTGGGATGAAAGGCGAGCGCCGAGGGGCAAGG--GG 3675  
Db 3244 GATCTGCTGAGCCCATGCGCATGCTGCTTGGCGAATATCATGTGTGGAATAATGGCCGC 3303  
Qy 3676 CACGATGGCTTTTACAGGCTCTCAGTACACCAAGACACCTTACGAGCGCCCTTAC 3735  
Db 3304 TTTTCTGGAATTCATCGACTGTGGCGGCTGGGTGGCGGACCGCTATCAGGACATAGCG 3363  
Qy 3736 ATGACAGCCCTGCGCCCTCGCTAAC-----TCGACGCGCGCGCGGATCCGGATTAAGTCCAA 3791  
Db 3364 TTGGCTACCGCTGATATTGCTGAAGAGCTTGGCGCGAATGGGCTGACCGCTTCTCTG 3423  
Qy 3792 TTTGTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTGTACTCAACAAATATCACAG 3851  
Db 3424 CTTTACGCTATCGCGCTCCCGATTCGACGCGCATCGCTTCTATCGCTTCTTGTGACGAG 3483  
Qy 3852 CTGAAGCCTATAGATACGAGCCATAGATAAAATAAAGATTTTATTTAGTCTCCAGAA 3911  
Db 3484 TTCTTCTGAGCGGACTCTGGGGTTCGATAAAATAAAGATTTTATTTAGTCTCCAGAA 3543



QY 3912 AAGGGGGGAATGAAAGACCCACCTGTAGGTTGGCAAGCTAGCTTAAAGTAAGCGCATTTT 3971  
DB |||||  
QY 3544 AAGGGGGGAATGAAAGACCCACCTGTAGGTTGGCAAGCTAGCTTAAAGTAAGCGCATTTT 3603  
DB |||||  
QY 3972 TGCAGGCGATGG-AAATAATACATACTGAGAAATAGAGAAAGTTAGATCAAGAGTTAGGAACA 4030  
DB |||||  
QY 3604 TGCAGGCGATGGAAGAAATACATACTGAGAAATAGAGAAAGTTAGATCAAGAGTTAGGAACA 3663  
DB |||||  
QY 4031 GA-GAGACAGCAGATATGGGCCCAACAGGATATCTGTGTAAGCAAGTTCTTCCGCC-CC 4088  
DB |||||  
QY 3664 GATGGAACAGCTGAATATGGGCCCAACAGGATATCTGTGTAAGCAAGTTCTTCCGCCGCCG 3723  
DB |||||  
QY 4089 TCAGGGCCCAAGAACAGTTGGAACAGGAGAAATATGGGCCCAACAGGATATCTGTGTAAGC 4148  
DB |||||  
QY 3724 TCAGGGCCCAAGAACAGATGGAACAGCTGAATATGGGCCCAACAGGATATCTGTGTAAGC 3783  
DB |||||  
QY 4149 AGTTCTGTCGCCGGCTCAGGGGCCAAGAACAGATATGTTCCAGATGCGGTCCGCCCTCAG 4208  
DB |||||  
QY 3784 AGTTCTGTCGCCGGCTCAGGGGCCAAGAACAGATATGTTCCAGATGCGGTCCAGGCCCTCAG 3843  
DB |||||  
QY 4209 CAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTCCCAAGGACCTGGAATGACCCCTGT 4268  
DB |||||  
QY 3844 CAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTCCCAAGGACCTGGAATGACCCCTGT 3903  
DB |||||  
QY 4269 GCCTTATTTGAACCTAACCAATCAGTTTCGCTTCTCGCTTCTGTTGGCGCTTCTGCTCC 4328  
DB |||||  
QY 3904 GCCTTATTTGAACCTAACCAATCAGTTTCGCTTCTCGCTTCTGTTGGCGCTTCTGCTCC 3963  
DB |||||  
QY 4329 CGAGCTCAATAAAGAGGCCCAACCCCTCACTCGGCGCGCAGTCTCCGATAGACTGC 4388  
DB |||||  
QY 3964 CGAGCTCAATAAAGAGGCCCAACCCCTCACTCGGCGCGCAGTCTCCGATGACTGA 4023  
DB |||||  
QY 4389 GTCCGCCGGGTACCGGTGTTCTCAATAAACCCCTCTGCAAGTTGCATCCGACTCGTGGTCT 4448  
DB |||||  
QY 4024 GTCCGCCGGGTACCGGTGTTCTCAATAAACCCCTCTGCAAGTTGCATCCGACTTGTGGTCT 4083  
DB |||||  
QY 4449 CGCTGTTCCTTGGAGGGTCT-CTCTGAGTGAATGACTACCCGTGAGCGGGTCTTTTCA 4507  
DB |||||  
QY 4084 CGCTGTTCCTTGGAGGGTCTCTCTGAGTGAATGACTACCCGTGAGCGGGTCTTTTCA 4143  
DB |||||  
QY 4508 TTTCTCCCACTACACAGCTCTCACTAACATTTCTGATGTCGCGCAGGACTCCGTGACG 4567  
DB |||||  
QY 4144 TTTGGGGGCTGTCGCGGATCGGAGACCCCTGCCAGGGAACACCGACCCACCGCGG 4203  
DB |||||  
QY 4568 CGGTTTTTTGTTTATAATAAATGCAAGAACAGTGTTCCTTCAAGCCAGACTACATCCT 4627  
DB |||||  
QY 4204 AGGTAAAGCTG-----GCT 4216  
DB |||||  
QY 4628 GACTCTCGGCTTTATAAAGAAATGTTGAAGGGTCTGTGGAGCTATCTGCCACACGACTTT 4687  
DB |||||  
QY 4217 GCCTCGCGGTTTCGGTGATGACGGTGAANAACCTCTG-ACATGCGAGTCCCGGAGCG 4275  
DB |||||  
QY 4688 TTAAGATTTTATGCTCCTGGATGAGGATTTAGTCAATCTATCTCTGCTCTATTTTGTCT 4747  
DB |||||  
QY 4276 GTCACAGCTTGTCTGTAAGCGGATGCGGAGCAGACAAGCCGTGAGGGCGGTGAGCG 4335  
DB |||||  
QY 4748 GGGTTCTCGTATTTAAATTTCTAGTTTGCACCTCCCTTCTGAGAGCAGCGGATGCA 4807  
DB |||||  
QY 4336 GGTGTTGGCG----- 4345  
DB |||||  
QY 4808 GAGTAGTTAATACTCTGAGGGCAGGCTTCTGTGAAAAGGTTGCTTGGGCTCAGTGTGAGA 4867  
DB |||||  
QY 4346 ----- 4345  
DB |||||  
QY 4868 TTTTGCCATAAAGGGGTCTGCCCCCTGTGTACAGACAGATCGGAATCTAGAGTGCATA 4927  
DB |||||  
QY 4346 ----- 4345  
DB |||||  
QY 4928 CTCAGAGTCCCGCGGTTTCGGGGCTCTGATCTCAGGGGATCTTTGCTCCTAGAGATCCTCT 4987  
DB |||||  
QY 4346 -----GCTGTGGGGCGCAGCCATGACCCAGTC----- 4373  
DB |||||  
QY 4988 ACGCCGACGATCGTGGCGGGGTACCGAGCTCGAATTCGTAATCATGTCATGCTGTT 5047

DB |||||  
QY 4374 -----ACGTAGCATAGCGAGTGTATCTGGCTTAACTATGCGCATCAGAGCAGAT 4426  
DB |||||  
QY 5048 TCCTGTGTGAATTTGTTATCCGCTCACAAATCCACACATACAGCGCGGAGCATAAA 5107  
DB |||||  
QY 4427 TGTACTGAGA----- 4436  
DB |||||  
QY 5108 GTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGGCTTGGCTCACT 5167  
DB |||||  
QY 4437 ----- 4436  
DB |||||  
QY 5168 GCCCGCTTTCCAGTCGGGAAACCTGTGTCGAGCTGCAATTAATGAATCGGCAACGCGC 5227  
DB |||||  
QY 4437 -----GTGCACCATATGCGGTGTGAAATACCCACAGATG 4471  
DB |||||  
QY 5228 GGGGAGAGCGGTTTGGCTATTTGGCGGCTCTTCCGCTTCTCGCTCACTGACTCGCTGGC 5287  
DB |||||  
QY 4472 CTTAAGGAGAAATATCCGCATCAGGCGCTCTTCCGCTTCTCGCTCACTGACTCGCTGGC 4531  
DB |||||  
QY 5288 CTGCGTCTGTCGCTGCGCGAGCGGTATCAGCTCACTCAAAAGGCGGTAAATACGGTTATC 5347  
DB |||||  
QY 4532 CTCGCTGCTTCCGCTGCGCGAGCGGTATCAGCTCACTCAAAAGGCGGTAAATACGGTTATC 4591  
DB |||||  
QY 5348 CACAGAATCAGGGGATAACGCAAGGAAAGAACATGTGAGCAAAAGGCCAGAAAAGCCAG 5407  
DB |||||  
QY 4592 CACAGAATCAGGGGATAACGCAAGGAAAGAACATGTGAGCAAAAGGCCAGAAAAGCCAG 4651  
DB |||||  
QY 5408 GAAACGTTAAAGGCGGCTTGTGCGGCTTTTCCATAGGCTCCGCGCCCTTACGAGCA 5467  
DB |||||  
QY 4652 GAAACGTTAAAGGCGGCTTGTGCGGCTTTTCCATAGGCTCCGCGCCCTTACGAGCA 4711  
DB |||||  
QY 5468 TCACAAAATCGACGCTCAAGTCAGAGGTGGGAAACCCGACAGACTATATAAGATACCA 5527  
DB |||||  
QY 4712 TCACAAAATCGACGCTCAAGTCAGAGGTGGGAAACCCGACAGACTATATAAGATACCA 4771  
DB |||||  
QY 5528 GCGCTTTCCCTCGGAAGCTCCCTGCTGCGCTCTCTGTTCCGACCCCTCCGCTTACCGG 5587  
DB |||||  
QY 4772 GCGCTTTCCCTCGGAAGCTCCCTGCTGCGCTCTCTGTTCCGACCCCTTACCGG 4831  
DB |||||  
QY 5588 ATACCTGTCGCGCTTTCTCCCTTCGGGAAAGCGTGGCGCTTCTCATAGCTACGCTGTAG 5647  
DB |||||  
QY 4832 ATACCTGTCGCGCTTTCTCCCTTCGGGAAAGCGTGGCGCTTCTCATAGCTACGCTGTAG 4891  
DB |||||  
QY 5648 GTATCTCAGTTCCGCTGAGTGGTTCGCTCCTCAAGCTGGCTGTGTCAGAACCCCGCT 5707  
DB |||||  
QY 4892 GTATCTCAGTTCCGCTGAGTGGTTCGCTCCTCAAGCTGGCTGTGTCAGAACCCCGCT 4951  
DB |||||  
QY 5708 TCAGCCCGACCGCTCGCTTATCCGCTTAACTATCGTCTTGAGTCCAAACCCCGTAAAGCA 5767  
DB |||||  
QY 4952 TCAGCCCGACCGCTCGCTTATCCGCTTAACTATCGTCTTGAGTCCAAACCCCGTAAAGCA 5011  
DB |||||  
QY 5768 CGACTTATCGCCACTCGGACGAGCCACTGGTAAACAGGATAGCAGAGCGAGTATGTAGG 5827  
DB |||||  
QY 5012 CGACTTATCGCCACTCGGACGAGCCACTGGTAAACAGGATAGCAGAGCGAGTATGTAGG 5071  
DB |||||  
QY 5828 CGGCTGCTACAGATTTCTTGAAGTGGTGGCTTAACTACGCTACACTAGAGGACAGTAT 5887  
DB |||||  
QY 5072 CGGCTGCTACAGATTTCTTGAAGTGGTGGCTTAACTACGCTTAACTAGAGGACAGTAT 5131  
DB |||||  
QY 5888 TGGTATCTGCGCTCTGCTGAAGCCAGTTTACCTTCGGGAAAGAGTTGGTAGTCTTGTATC 5947  
DB |||||  
QY 5132 TGGTATCTGCGCTCTGCTGAAGCCAGTTTACCTTCGGGAAAGAGTTGGTAGTCTTGTATC 5191  
DB |||||  
QY 5948 CGGCAAAACAAACCCAGCTGGTAGCGGTGTTTTTTTTTGTTCGACAGCAGATTAACGGG 6007  
DB |||||  
QY 5192 CGGCAAAACAAACCCAGCTGGTAGCGGTGTTTTTTTTTGTTCGACAGCAGATTAACGGG 5251  
DB |||||  
QY 6008 CAGAAAAGAGATCTCAAGAGATCTTTGATCTTTTCTACGGGTCTGACGCTCAGTG 6067  
DB |||||  
QY 5252 CAGAAAAGAGATCTCAAGAGATCTTTGATCTTTTCTACGGGTCTGACGCTCAGTG 5311  
DB |||||  
QY 6068 GAAACGAAATCTCAGCTTAAAGGATTTTGTTCATGAGATTTTCAAAAAGGATCTTCACTTA 6127

Db 5312 GAACGAAACCTCACGTTAAGGATTTTGGTCATGAGATTTATCAAAAAGGATCTTCACCTA 5371  
Qy 6128 GATCCTTTTAAATTAATAAATGAAGTTTAAATCAATCTAAAGTATATATGAGTAAACTTGG 6187  
Db 5372 GATCCTTTTAAATTAATAAATGAAGTTTAAATCAATCTAAAGTATATATGAGTAAACTTGG 5431  
Qy 6188 GTCTGACAGTTACCAATGCTTAATCAAGTGAAGGACCTATCTCAGGATCTCTCTATTTCG 6247  
Db 5432 GTCTGACAGTTACCAATGCTTAATCAAGTGAAGGACCTATCTCAGGATCTCTCTATTTCG 5491  
Qy 6248 TTCAATCCATAGTTGCTGACTCCCGTCTGTAGATACTACGATACGAGGAGGCTTACC 6307  
Db 5492 TTCAATCCATAGTTGCTGACTCCCGTCTGTAGATACTACGATACGAGGAGGCTTACC 5551  
Qy 6308 ATCTGGCCCCAGTGTGCAATGATACCGCGAGACCCAGCTCACGGGCTCCAGATTATTC 6367  
Db 5552 ATCTGGCCCCAGTGTGCAATGATACCGCGAGACCCAGCTCACGGGCTCCAGATTATTC 5611  
Qy 6368 AGCAATAACACGACCGGAGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6427  
Db 5612 AGCAATAACACGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5671  
Qy 6428 CTCAATCCAGTCTATTAAATGTTGCGGGAAGCTAGAGTAAGTATGTCGCGAGTTAATAG 6487  
Db 5672 CTCAATCCAGTCTATTAAATGTTGCGGGAAGCTAGAGTAAGTATGTCGCGAGTTAATAG 5731  
Qy 6488 TTTGGGCAACGTTGTTGCCATGCTGCTACAGGC-TGCTGGTGTGACGCTCGTCTGTTGGTAT 6546  
Db 5732 TTTGGGCAACGTTGTTGCCATGCTGCTACAGGCATGCTGGTGTGACGCTCGTCTGTTGGTAT 5791  
Qy 6547 GGGTTCATAGCTCCGGTTCCTCAACGATCAAGGGAGTTACATGATCCCCCATGTTGTG 6606  
Db 5792 GGGTTCATAGCTCCGGTTCCTCAACGATCAAGGGAGTTACATGATCCCCCATGTTGTG 5851  
Qy 6607 CAAAAAGCGGTAGCTCTCGTCTCGTCCGATCGTGTGTCAGAGTAAAGTGGCGCGAGT 6666  
Db 5852 CAAAAAGCGGTAGCTCTCGTCTCGTCCGATCGTGTGTCAGAGTAAAGTGGCGCGAGT 5911  
Qy 6667 GTTATCACTCATGTTATGCGCAGCACTGCATAAATTTCTTACTGTATGCCATCCGTAAG 6726  
Db 5912 GTTATCACTCATGTTATGCGCAGCACTGCATAAATTTCTTACTGTATGCCATCCGTAAG 5971  
Qy 6727 ATGCTTTTCTGTGAGTGTGAGTACTCAACCAAGTCAATCTGAGAATAGTGTATGCGGCG 6786  
Db 5972 ATGCTTTTCTGTGAGTGTGAGTACTCAACCAAGTCAATCTGAGAATAGTGTATGCGGCG 6031  
Qy 6787 ACCGAGTTGCTTTCGCGCGGTCAATACGGGTAATACCGGCAATACGAGCACTTT 6846  
Db 6032 ACCGAGTTGCTTTCGCGCGGTCAACACGGGTAATACCGGCAATACGAGCACTTT 6091  
Qy 6847 AAAAGTGTCTCATTTGGAACACGTTCTTTCGGGCGAAAACTCTCAAGGATCTTACCGCT 6906  
Db 6092 AAAAGTGTCTCATTTGGAACACGTTCTTTCGGGCGAAAACTCTCAAGGATCTTACCGCT 6151  
Qy 6907 GTTGAGATCCAGTTGATGTAACCACTCGTGTGCAACCACTGATCTTCAGCATCTTTTAC 6966  
Db 6152 GTTGAGATCCAGTTGATGTAACCACTCGTGTGCAACCACTGATCTTCAGCATCTTTTAC 6211  
Qy 6967 TTTTACCAGCGTTTCTGGGTGAGCAAAACAGGAGGCAAAATCCGCAAAAAGGGAAT 7026  
Db 6212 TTTTACCAGCGTTTCTGGGTGAGCAAAACAGGAGGCAAAATCCGCAAAAAGGGAAT 6271  
Qy 7027 AAGGCGCAGACGGAATGTTGAAATCTCATACTCTTCTCTTTTCAATATTATGAAGCAT 7086  
Db 6272 AAGGCGCAGACGGAATGTTGAAATCTCATACTCTTCTCTTTTCAATATTATGAAGCAT 6331  
Qy 7087 TTATCAGGGTTATTGTTCTCATGAGCGGATACATATTTGAATGATATTTAGAAAAATAACA 7146  
Db 6332 TTATCAGGGTTATTGTTCTCATGAGCGGATACATATTTGAATGATATTTAGAAAAATAACA 6391  
Qy 7147 AATAGGGGTTCCGGGCACTTTCCCGGAAAAAGTGGCACCTGACGCTTAAGAAACCATTAT 7206  
Db 6392 AATAGGGGTTCCGGGCACTTTCCCGGAAAAAGTGGCACCTGACGCTTAAGAAACCATTAT 6451

Qy 7207 TATCATGACATTAACCTATATAAATAGCGTATCAAGAGGCCCTTTCCTCT 7257  
Db 6452 TATCATGACATTAACCTATATAAATAGCGTATCAAGAGGCCCTTTCCTCT 6502  
RESULT 14  
BD138563  
LOCUS  
DEFINITION BD138563 6522 bp DNA linear PAT 18-SEP-2002  
Expression system of nucleotide with reduced immunogenicity for use  
in gene therapy.  
ACCESSION BD138563  
VERSION BD138563.1 GI:232333508  
KEYWORDS JP 2002508976-A/2.  
SOURCE synthetic construct  
ORGANISM synthetic construct  
other sequences; artificial sequences.  
REFERENCE 1 (bases 1 to 6522)  
Radosevich,T.J. and Jr,C.J.L.  
AUTHORS Expression system of nucleotide with reduced immunogenicity for use  
in gene therapy  
TITLE Patent: JP 2002508976-A 2 26-MAR-2002;  
JOURNAL HUMAN GENE THERAPY RESEARCH INSTITUTE  
COMMENT OS Artificial Sequence  
PN JP 2002508976-A/2  
PD 26-MAR-2002  
PF 13-JAN-1999 JP 2000540263  
PR 14-JAN-1998 US 60/071409  
PI THOMAS J RADOSEVICH CHARLES J LINK JR  
PC C12N15/09.A61K31/711.A61K48/00.A61P37/06.C12N1/15.C12N1/19, PC  
C12N1/21,  
PC C12N5/10/A61K35/76, (C12N15/09, C12R1:93), C12N15/00, C12N5/00,  
PC (C12N15/00, C12R1:93)  
CC Description of Artificial Sequence : recombinant vector FH  
Key Location/Qualifiers  
FT source 1..6522  
/organism='Artificial Sequence'.  
FEATURES  
source  
1..6522  
/organism='synthetic construct'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:32630'  
ORIGIN  
Query Match 40.4%; Score 3096; DB 6; Length 6522;  
Best Local Similarity 70.8%; Pred. No. 0;  
Matches 4803; Conservative 0; Mismatches 1465; Indels 518; Gaps 31;  
Qy 549 TTTTGAAGAGCCACCTGTAGGTTTGGCAAGCTTAGCTTAAGTAAACGCAATTTTGAAGG 608  
Db 174 TTTTGAAGAGCCACCCACCGTAGG--TGGCAAGCTAGCTTAAGTAAACGCACTTTGCAAGG 231  
Qy 609 CATGG-AAAATACATACCTGAGATAGAGATTTCAGATCAAGTTAGGACACAGAGAC 667  
Db 232 CATGGAAAAATACATACCTGAGAAATGAAAAAGTTTCAGATCAAGTTCAGGAAACAAGAAAC 291  
Qy 668 AGCAAGATATGGGCAACCAACAGGATATCTGTGTGAAGCAGTTCCTCGCCCC--GCTCAGGGGCC 726  
Db 292 AGCTGAATA---CCAAACAGGATATCTGTGTGAAGCGGTTCCTGCCCCGCTCAGGGCC 347  
Qy 727 AAGAACAGTTGGAAACAGAGAAAT-ATCGGCGCAACAGAGATATCTGTGTGAAGCAGTTCCT 785  
Db 348 AAGAACAGATGAGACAGCTGAGTGTGATGGCCAAACAGAGATATCTGTGTGAAGCAGTTCCT 407  
Qy 786 GCCCCGGCTCAGGCCCAAGAACAGATGTTCCCGAGATGCGGTCCCGCCCTCAGCAGTTTC 845  
Db 408 GCCCCGGCTCAGGCCCAAGAACAGATGTTCCCGAGATGCGGTCCCGCCCTCAGCAGTTTC 467  
Qy 846 TAGAGAACCATTCAGATGTTTCCAGGGTCCCGCAAGGACCTG-AAATGACCTGTGCGCTTA 904  
Db 468 TAGTGAATCATCAGATGTTTCCAGGGTCCCGCAAGGACCTGAAATGACCTGTGACCTTA 527  
Qy 905 TTTGAACCAATCATGTTTCGCTTCTTCGCTTCTGTGTCGCGCGCTTCTGTGTCGCCGAGCT 964

|||||  
528 TTGAACTAACCAATCAGTTTCGCTTCTCGCTTCGCGCGCTTCGCTCCGAGCT 587  
QY  
965 CAATAAAGAGCCCAACACCCCTCACTCGCGCGCCAGTCTCCGATAGACTGCGTCGCC 1024  
DB  
588 CAATAAAGAGCCCAACACCCCTCACTCGCGCGCCAGTCTTCGATAGACTGCGTCGCC 647  
QY  
1025 CGGGTACCGGTATTTCCCAATAAAGCCTCTTGTCTTTGATCGGAATCGTGACTCGCTG 1084  
DB  
648 CGGGTACCGGTATTTCCCAATAAAGCCTCTTGTCTTTGATCGGAATCGTGACTCGCTG 707  
QY  
1085 ATCTCTGGGAGGCTCTCTCAGATTGATGACTGCCCCACTCGGGGCTCTTTCAATTGGA 1144  
DB  
708 TTCTCTGGGAGGCTCTCTCAGTGTGATGACTACCCACGACGGGGTCTTTCAATTGGG 767  
QY  
1145 GGTTCACACGAGATTGGAGACCCCTGCCAGGGACCAACGACCCCGCGCGGAGGTA 1204  
DB  
768 GGCTCTCGGGATTGGAGACCCCTGCCAGGGACCAACGACCCCGCGGAGGTA 826  
QY  
1205 AGCTGSCCAGCACTTATCTGTCTCTGCTCGGATTGCTAGTGTCTATGACTGATTTATG 1264  
DB  
827 AGCTGSCCAGCACTTATCTGTCTCTGCTCGGATTGCTAGTGTCTATGATTGATTG 886  
QY  
1265 CGCTCTCGCTCGGTACTAGTTAGTAACTAGTCTGTATCTGCGGACCCGCTGGTGAAC 1324  
DB  
887 CGCTCTCGCTCGGTACTAGTTAGTAACTAGTCTGTATCTGCGGACCCGCTGGTGAAC 946  
QY  
1325 GAGGATTCGGAACACCCCGCGCAACCCCTGGGAGACGTCGCAGGGACTTCGGGGCCGT 1384  
DB  
947 GAGGATTCGGAACACCCCGCGCAACCCCTGGGAGACGTCGCAGGGACTTCGGGGCCGT 1006  
QY  
1385 TTTTGTGGCCGACCTGAGTCTTAATAATCCCGATCGTTTAGGACTCTTTTGGTGCACCCC 1444  
DB  
1007 TTTTGTGGCCGACCTGAGGAAAGGAGTCAATGTGGAATCCGACCCCGTCA 1057  
QY  
1445 CTTAGAGGAGGATATGTGTTCTGTTAGGAGACGAGAACCTTAAACAGTTCCCGCTCC 1504  
DB  
1058 -----GGATATGTGTTCTGTTAGGAGACGAGAACCTTAAACAGTTCCCGCTCC 1107  
QY  
1505 GTCTGAATTTTGTCTTCGTTTGGGACGAGAGCCGCGCGCTGTTGT---CTGCTG 1561  
DB  
1108 GTCTGAATTTTGTCTTCGTTTGGAAACGAGCCGCGCGCTGTTGTCTGCTGCGCGCTG 1167  
QY  
1562 CAGCATCGTCTGTGTTGCTCTGCTGACTGTGTTTCTGTTATGTTCTGAAATATGG 1621  
DB  
1168 CAGCATCGTCTGTGTTGCTCTGCTGACTGTGTTTCTGTTATGTTCTGAAATATGG 1227  
QY  
1622 CCGGGCTAGACTGTTTACCACTCCCTTAAAGTTTGAACCTTAGGTCACTGGAAAGATGCGA 1681  
DB  
1228 CC-----AGACTGTTTACCACTCCCTTAAAGTTTGAACCTTAGGTCACTGGAAAGATGCGA 1281  
QY  
1682 GCGGATCGCTCAACACAGTCGGTAGATGTCAAGAAAGAGAGCGTTGGGTACCTTCTGCTC 1741  
DB  
1282 GCGGATCGCTCAACACAGTCGGTAGATGTCAAGAAAGAGAGCGTTGGGTACCTTCTGCTC 1341  
QY  
1742 TGCAGAAATGCCAAACCTTTAAAGTCGGATGGCGGAGCGGACCTTTAAACGAGACCT 1801  
DB  
1342 TGCAGAAATGCCAAACCTTTAAAGTCGGATGGCGGAGCGGACCTTTAAACGAGACCT 1401  
QY  
1802 CATCACCCAGGTTAAGATCAAGGTCTTTTACCTGGCCGCGATGGACACCCAGACAGGT 1861  
DB  
1402 CATCACCCAGGTTAAGATCAAGGTCTTTTACCTGGCCGCGATGGACACCCAGACAGGT 1461  
QY  
1862 CCCCTACATGCTGACCTGGGAAGCCTTGGCTTTTGAACCCCTCCCTGGGTCAAGCCCTT 1921  
DB  
1462 CCCCTACATGCTGACCTGGGAAGCCTTGGCTTTTGAACCCCTCCCTGGGTCAAGCCCTT 1521  
QY  
1922 TGTACACCTTAAGCCTCCGCTCTCTTCTCCATCCGCGCGCTCTCCCTTGAACC 1981  
DB  
1522 TGTACACCTTAAGCCTCCGCTCTCTTCTCCATCCGCGCGCTCTCTCCCTTGAACC 1581  
QY  
1982 TCTCTGTTGACCCCGCTCGATCTCTCTCTTTATCCAGGCCCTCACTCTTCTCTAGGCGC 2041

DB  
1582 TCCTCGTTTCGACCCCGCTCGATCTCTCCCTTTATCCAGCCCTCACTCTTCTCTAGGCGC 1641  
QY  
2042 CCCCATATGGCCATATGAGATCTTATATGGGGCACCCCGCCCTTGTAAATCTCCCTGA 2101  
DB  
1642 CGGATGAACCTTGTATATGTTTCTA---GCCCTCTGGGGCCCGGTTCGGGGTAGTAT 1697  
QY  
2102 CCTTGACATGACAGAGTTTACTTAAACAGCCCTCTCTTCCAAAGCTCACTTTACAGGCTTTTAC 2161  
DB  
1698 GCTCGAATATCTCTTGACTCTTTTCGATGAACTCTCCGCCCTTGGTGAGACGAGCGGTT 1757  
QY  
2162 TTATGTCAGACAGAAAGTCTGGAGACCTCTGGCGGAGCCTACCAAGAAACAATGGACCGA 2221  
DB  
1758 ACCGCTCTGTCCGATGTTTCGAGTACCGAGTAGAGTATTCGAGGCGCGTTCGGTGCT 1817  
QY  
2222 CGGTTGTTACCTCAACCTTACCGAGTCGGGACACAGTGTGGTCCGCCGACACAGACT 2281  
DB  
1818 CCGATCGGGCGGTGCACTGGAGGCTCTGTGAAACCTTGGCGGGAACC-----TGT 1867  
QY  
2282 AAGAAACCTAGAAACCTCGCTGGAAAGGACCTTACACAGTCTCTGCTGACCAACCCGCCCGCC 2341  
DB  
1868 CCGTGCCACGCGGACACCCCGGTTGTACTACAGACGCTGGAGGCTACGGGATCGAG 1927  
QY  
2342 CTCAAAGTAGACGGCATCGAGCTTGGATACACGGCGGCCACGTGAAGGCTGCGGACCCC 2401  
DB  
1928 TGCCGACGCGGTTGGAGGACGCTCCGAAAGCCTCGTCGAAACGCTACTTGGCTCCGG 1987  
QY  
2402 GGGGTTGNAACCTCTTAGACTGCCATCGGATGGAGCTGATCATCTCTCTTCTTGGTAG 2461  
DB  
1988 ACTATCGTGTTCCTCCAAACGCAACAACTCGTGTGTCTTCTTCCCTCCGCCAAT 2047  
QY  
2462 CAACAGCTACAGTGTCCACTCCGACTCCAGCTACCCAGAGCCCAAGCAGCTGAGCG 2521  
DB  
2048 GCCAACTTATATGTAGAGTCGAAACCCCGGTGCTCGTCTTGGGTTCCCTGTGGA 2107  
QY  
2522 CCAGCGTGGGTGACAGAGTGACCATCACTGTAAAGCCAGTCAGAGTGGGTACTT--- 2578  
DB  
2108 GCTCGTTAGAGGACATCGAAACGACTATTGTTGAAAGTCGCGCTCTAATGGGCTACTAG 2167  
QY  
2579 -----CTGTAGCTTGGTACACAGCAGAGCAGGTAAGCTCCAAAGCTGCTACTACT 2632  
DB  
2168 CGCTCAGCATTAAGTCGGCGCAGTATACGCTGATGTTGGCAGTGAATTCAGTGTTT 2227  
QY  
2633 GGACATCCACCCCGCACACTGTTGTGCCAAGCAGATTACGCGGTAGCGGTAGCGGTACCG 2692  
DB  
2228 GGGGCTGTATGTGAAGGTTGGCTG-CACCCGACATTTTCCCTGGATGTTTTCGACAG 2286  
QY  
2693 ACTTCACCTTACCATCAGACGCTCCAGCCAGAG-----GACATCGCCACCTA 2741  
DB  
2287 TGGTGAAATTCGTTAACTCGAGGATCCGGCTGTGGAATGTGTGAGTTAGGGTGTGGA 2346  
QY  
2742 CTACTGCCAGCAATATAGCTCTATCGGTCTGCGCCAGGGACCAAGGTGGAATCAA 2801  
DB  
2347 AGTCCCGAGGCTCCCGCAGCAGGCAAGATGCAAAAGCATGCTCTCAATAGTCAGCAA 2406  
QY  
2802 ACGAGTGGCTCAGGATCGGGTGGATCCGGCTCTGTGTGCTCAGGATCGGAGTCCCACT 2861  
DB  
2407 CAGGTGTGAAAGTCCCGAGGCTCCCGAGGAGGAGATGCAAGCATGCACTCA 2466  
QY  
2862 GGTGAGAGCGGTGAGGTGTTGTGCAACCTCGCGCGGTCCCTGCGCTCTGTCTGCTCCG 2921  
DB  
2467 ATTAGTCAGCAACCATAGTCCCGCCCTTAACTCCGCCCATCCCGCCCTTAACTCCGCCCA 2526  
QY  
2922 ATCTGGCTCGATTTTCCACATATTTGGATGAGTTGGGTGAGACAGGACCTCGGAAAGG 2981  
DB  
2527 GTTCGCGCCCATTTCTCCGCCCATGCTGACTAAATTTTTTTTATTTATGACAGAGCGGAG 2586  
QY  
2982 TCTTGAGTGGATTGGAGAAATTCATCCAGATAGACTAGATTAATCTATGCGCGCTCTCT 3041  
DB  
2587 CGGCTCGGCTCTGAGCTATTTCCAGAGTAGTGAGGAGCTTTTTTGGAGGCTTAGGCT 2646  
QY  
3042 AAAGGATAGATTTA---CAATATCGGAGACAAAGCAAGAAACATATTTCTCTGCAAT 3098  
DB  
2647 TTTGCAAAAAGCTTGGGCTGCGAGGTCGAGGCGGATCTGATCAAGAGACAGGATGAGGATC 2706

QY 3099 GGACAGCCTGAGACCCGGAAGACACCGGGTCTATTTTGTGAGCCCTTTACTCGGGTT 3158  
Db |||||  
QY 2707 GTTTCGATGATGAAAGATGGAATGCACGAGGTTCTCGGGCGCTTGGGTGAGAG 2766  
Db |||||  
QY 3159 CCCCTGGTTTGTCTATTGGGGCCAGGACCCCGGTCAACCGTCTCCAGTGTCTAAGCCAC 3218  
Db |||||  
QY 2767 GCTATTGGCTATGACTGGGCAACAAGACAAATCGGCTGCTCTGATGCCGCGTGTCCG 2826  
Db |||||  
QY 3219 CACGACGCGGCGGACCAACCAACACCGCGGCCACCATCGCGTCCGAGCCCTGTC 3278  
Db |||||  
QY 2827 GCTGTACGCGGCGCGCGTCTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAA 2886  
Db |||||  
QY 3279 CTTGGCCCAAGGCGG-----CTCGGCACGCGGCGGGGCGGAGTGCACAC 3326  
Db |||||  
QY 2887 TGAATCTGAGGACGAGGCGGCGCTATCGTGGCTGGCCACGACGCGGCTTCTTGGCG 2946  
Db |||||  
QY 3327 GAGGGGCTGACTTTCGCTGGATCCAAACTCTGTACTCTGTGGATGGAATCTCTT 3386  
Db |||||  
QY 2947 AGCTGTGCTCGACGTTGTCACTGAAGCGGAGGACTGGCTGTCTATTGGGCGAGTGCC 3006  
Db |||||  
QY 3387 CATCTAATGGTGTCAATCTCACTGCGCTTCTCTGAGAGTGAAGTTTCAGCAGGAGCGCAGA 3446  
Db |||||  
QY 3007 GGGCAGGATCTCTCTGTCTATCTCACTTGTCTCTGCGAGAAAGTATCCATCATGGCTGA 3066  
Db |||||  
QY 3447 GCC-----CCCCGGGTACCAGCAGGCGCCAGAACCCAGCTCTA 3482  
Db |||||  
QY 3067 TGAATCGGGGGTGCATAGCTGTGATCCGGCTACCTGCGCAATTCGACACCAAGCGAA 3126  
Db |||||  
QY 3483 TAACGAGCTCAATCTAGGACGAGAGGATGATGATGTTTGGACAGAGACGTGGCG 3542  
Db |||||  
QY 3127 ACATCGCATCAGCGAGCACATCTCGGATGGAACCGGCTTGTGATCAGGATGATCT 3186  
Db |||||  
QY 3543 GGACCTTGAGATGGGGGAAAGCCGAGAAAGAAACCCCTCAGGAAGCCCTGTACAATGA 3602  
Db |||||  
QY 3187 GGACGAGAGCATCAGGG--GCTCGGCGCAGCGAATGTTGCGCAAGGCTCAAGGCGCGC 3244  
Db |||||  
QY 3603 ACTCGAAGAGATAGATGGCGAGGCTCAAGTGAAGTGGGATGAAAGCGGAGCGCGC 3662  
Db |||||  
QY 3245 ATGCCCGACGCGAGGATCTGCTGTCACCATGCGCATGCGCTGCTGCGCAATCATG 3304  
Db |||||  
QY 3663 GAGGGCAGG--GGCAGATGGCTTTACAGGGTCTCAGTACAGCCACCAAGACAC 3720  
Db |||||  
QY 3305 GTGAAATGGCCGCTTTTCTGGATTCATGACTGTGGCGGGCTGGGTGGCGGAGCGC 3364  
Db |||||  
QY 3721 TACGACGCTTCATATGACGCGCTCGCCCTCGCTAAC-----TCGACGCGCGCGGA 3776  
Db |||||  
QY 3365 TATCAGNATAGCGTTGGCTACCGTGATATGCTGAGAGCTTGGCGGCAATGGCT 3424  
Db |||||  
QY 3777 TCCGATPAGTCCAAATTTGTTAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTGA 3836  
Db |||||  
QY 3425 GACCGCTTCTCGTGTCTTACGGTATCGCGCTCCCGATTCGAGCGCATCGCTTCTAT 3484  
Db |||||  
QY 3837 CAACAATATCAGCTGAGCCCTATAGATGACGACCATAGATAAATAAAGATTTTA 3896  
Db |||||  
QY 3485 GCGCTTCTTGACGAGTCTTCTGACGCGGACTCTGGGGTTCGATAAATAAAGATTTTA 3544  
Db |||||  
QY 3897 TTTAGTCTCCAGAAAAGGGGGGAAATGAAAGACCCACCTCTAGTGTGGCAAGCTAGCT 3956  
Db |||||  
QY 3545 TTTAGTCTCCAGAAAAGGGGGGAAATGAAAGACCCACCTCTAGTGTGGCAAGCTAGCT 3604  
Db |||||  
QY 3957 TAAGTAACGCCATTTTGAAGGCTAGG--AAATAACATAACTGAGAATAGAGAAGTTTCA 4015  
Db |||||  
QY 3605 TAAGTAACGCCATTTTGAAGGCTAGGAAATAATACATAACTGAGAATAGAGAAGTTTCA 3664  
Db |||||  
QY 4016 TCAAGTTAGGAACAGA--GAGACAGCAGATATGGGCCAAACAGAGATATCTGTGGTAAGC 4074  
Db |||||  
QY 3665 TCAAGGTGAGGAACAGATGGAACAGCTGAATATGGGCCAAACAGAGATATCTGTGGTAAGC 3724  
Db |||||  
QY 4075 AGTTCTTCCGCC--GCTCAGGGCCAGAACAGTGTGMAACAGGAGATATGGGCCAAACAGG 4133  
Db |||||  
QY 3725 AGTTCTTCCGCCCTTCHGGGCCAAGAACAGATGGAACAGCTGAATATGGGCCAAACAGG 3784  
Db |||||

QY 4134 ATATCTGTGTAGAGAGTTCTGCCCCGGCTCAGGGCCCAAGAACAGATGGTCCCAGATG 4193  
Db |||||  
QY 3785 ATATCTGTGTAGAGAGTTCTGCCCCGGCTCAGGGCCCAAGAACAGATGGTCCCAGATG 3844  
Db |||||  
QY 4194 CGGTCCGCCCTCAGCAGTTTCTAGAGAAACATCAGATGTTTCCAGGGTGCCCAAGGAC 4253  
Db |||||  
QY 3845 CGGTCCAGCCCTCAGCAGTTTCTAGAGAAACATCAGATGTTTCCAGGGTGCCCAAGGAC 3904  
Db |||||  
QY 4254 CTGAAATGACCTGTGCTTATTTTGAACCTAACCAATCAGTTCTCGTTCCTGTTTCG 4313  
Db |||||  
QY 3905 CTGAAATGACCTGTGCTTATTTTGAACCTAACCAATCAGTTCTCGTTCCTGTTTCG 3964  
Db |||||  
QY 4314 CGCGTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCAGTCCGGGCGCCAGT 4373  
Db |||||  
QY 3965 CGCGTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCAGTCCGGGCGCCAGT 4024  
Db |||||  
QY 4374 CCTCGATAGACTCGCTCGCCGGGTACCCGTGTTCTCAATAAACCCCTCTTGCAGTTGCA 4433  
Db |||||  
QY 4025 CCTCGATGACTGAGTCCCGGGTACCCGTGATCCAAATAAACCCCTCTTGCAGTTGCA 4084  
Db |||||  
QY 4434 TCCGACTCGTGGTCTCGCTGTTCTTGGGAGGGTCT--CTCTGAGTGAATGACTACCCGTC 4492  
Db |||||  
QY 4085 TCCGACTGTGCTCTCGCTGTTCTTGGAGGGTCTCTCTGAGTGAATGACTACCCGTC 4144  
Db |||||  
QY 4493 AGCGGGTCTTTTCAAGTTTCTCCCACTACAGGCTCTCACTAACTCTGATGTCGCGC 4552  
Db |||||  
QY 4145 AGCGGGGCTCTTCAATTTGGGGGCTCGTCCGGATCGGAGACCCCTGCCAGGAGCCAC 4204  
Db |||||  
QY 4553 AGGACTCCGTCAGCCGGTCTTGTATTAATAAATGCAAGACAGTGTCCCTTCA 4612  
Db |||||  
QY 4205 CGACCCACACCGGAGGTAAGCTG----- 4229  
Db |||||  
QY 4613 GCCAGACTACATCTCGCTCTTATAAAGAAATGTTGAAGGGCTCTGTGGACTAT 4672  
Db |||||  
QY 4230 -----GCTGCTCGCGGTTTCTGATGACGGTGAAACCTCTG--ACACATG 4276  
Db |||||  
QY 4673 CTGCCACAGACTTTTAAAGATTTTATGCTCTCTGATGAGGATTTAGTCAATCTATC 4732  
Db |||||  
QY 4277 CAGCTCCGGAGACGGTCAACAGTGTGTGTAAGCGGATGCGGGAGCAGACAACCCGT 4336  
Db |||||  
QY 4733 CTGCTATTTTGTGCTGCTTCTCGCTATTTTAAATTTCTAGTTTGCATCCCTCCCTGAG 4792  
Db |||||  
QY 4337 CAGGCGGCTCAGCGGGTGTGGCG----- 4361  
Db |||||  
QY 4793 AGCAGCGGATGACAGAGTAGTTAATACTCTGAGGCGAGGCTTCTGTGAAAAGGTTGCT 4852  
Db |||||  
QY 4362 ----- 4361  
Db |||||  
QY 4853 GGGCTCAGTGTGAGATTTTGCCATAAAGGGGCTCTGCCCTGTGTACACAGATCGG 4912  
Db |||||  
QY 4362 ----- 4361  
Db |||||  
QY 4913 AATCTAGAGTGCATCTCAGAGTCCCGGCTTCCGGGCTCTGATCTCAGGGCATCTTT 4972  
Db |||||  
QY 4362 -----GGTGGGGGCGCAGCCATGACCCAGTC--- 4389  
Db |||||  
QY 4973 GCCTAGAGATCTCTACGCGGAGCATCGTGGCGGGTACCGAGCTCGAATTCGTAATC 5032  
Db |||||  
QY 4390 -----ACGTAGCAGTACGGAGTGATATCTGGCTTAACATATGC 4427  
Db |||||  
QY 5033 ATGCTCATAGCTGTTTCTGTGGAATTTGTTATCCGCTCAAAATCCACAACATACG 5092  
Db |||||  
QY 4428 GGCAATCAGAGCAGATTGTACTGAGA----- 4452  
Db |||||  
QY 5093 AGCCGGAAGCAATAAGTGTAAAGCCTGGGTGCTTAATGAGTGAAGTCAACATTAAT 5152  
Db |||||  
QY 4453 ----- 4452  
Db |||||  
QY 5153 TGGTGGCTCACTGCCCGCTTTCAGTCCGGAACCTGTGTCGCGCAGCTGCAATTAATG 5212  
Db |||||  
QY 4453 -----GTGACCAATATGCGGTGGA 4472  
Db |||||  
QY 5213 AATCGGCAACGCGCGGGAGAGCGGTTTGCATTATTTGGGCGCTTTCGCTTCTCGCT 5272  
Db |||||

Db 4473 AATACCGCAGATGCTTAAGAGAGAAAATACCGCATCAGCGCGCTCTTCGCTCTCGCT 4532  
 QY 5273 CACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 5332  
 Db 4533 CACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 4592  
 QY 5333 GGTAAATACGGTTATCCACAGAAATCAGGGATACCGAGGAAAGAACATGTGAGCAAAAGG 5392  
 Db 4593 GGTAAATACGGTTATCCACAGAAATCAGGGATACCGAGGAAAGAACATGTGAGCAAAAGG 4652  
 QY 5393 CCAGCAAAAGGCGAGGAAACCGTAAAGGCGGTTGCTGGGCTTTTCCATAGGCTCG 5452  
 Db 4653 CCAGCAAAAGGCGAGGAAACCGTAAAGGCGGTTGCTGGGCTTTTCCATAGGCTCG 4712  
 QY 5453 CCCCTGACGAGCATCAAAAAATCGAGCTCAAGTCAAGAGTGGCGAAACCCGACAGG 5512  
 Db 4713 CCCCTGACGAGCATCAAAAAATCGAGCTCAAGTCAAGAGTGGCGAAACCCGACAGG 4772  
 QY 5513 ACTATAAGATACCGAGGCTTTCCCTCGAAGCTCCCTCGTGGCTCTCTGTTCCGAC 5572  
 Db 4773 ACTATAAGATACCGAGGCTTTCCCTCGAAGCTCCCTCGTGGCTCTCTGTTCCGAC 4832  
 QY 5573 CTTGCGGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCTTTCTCA 5632  
 Db 4833 CTTGCGGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCTTTCTCA 4892  
 QY 5633 TAGCTACCGCTGAGGTATCTCAGTTCCGTTGAGTGGTGGTCCGTCAGCTGGCTGGTGT 5692  
 Db 4893 TAGCTACCGCTGAGGTATCTCAGTTCCGTTGAGTGGTGGTCCGTCAGCTGGCTGGTGT 4952  
 QY 5693 GACGAAACCCCGTTACCGCGAGCGCTGCGCTTTATCCGTTACTATCTGTTGAGTC 5752  
 Db 4953 GACGAAACCCCGTTACCGCGAGCGCTGCGCTTTATCCGTTACTATCTGTTGAGTC 5012  
 QY 5753 CAACCGGTAAGACAGCACTTATCGCACTGGCAGCAGCACTGGTAAACAGGATTAGCAG 5812  
 Db 5013 CAACCGGTAAGACAGCACTTATCGCACTGGCAGCAGCACTGGTAAACAGGATTAGCAG 5072  
 QY 5813 AGGAGGTATGAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACTACGGCTACAC 5872  
 Db 5073 AGGAGGTATGAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACTACGGCTACAC 5132  
 QY 5873 TAGAAGCAGATTTTGGTATCTGGCTCTGCTGAGCGAGTTACCTCGGAAAGAGCT 5932  
 Db 5133 TAGAAGCAGATTTTGGTATCTGGCTCTGCTGAGCGAGTTACCTCGGAAAGAGCT 5192  
 QY 5933 TGGTAGCTCTTGATCCGGCAAAACAAACCGCTGGTAGCGGTGGTTTTTTTGTGCAA 5992  
 Db 5193 TGGTAGCTCTTGATCCGGCAAAACAAACCGCTGGTAGCGGTGGTTTTTTTGTGCAA 5252  
 QY 5993 GCAGCAGATTAACCGCAGAAAAAAGGATCTCAAGAGATCTTTTGTATCTTTTACGGG 6052  
 Db 5253 GCAGCAGATTAACCGCAGAAAAAAGGATCTCAAGAGATCTTTTGTATCTTTTACGGG 5312  
 QY 6053 GTCTGACGCTCAGTGAAACGAAACCTCAGTTAAGGATTTTGGTCAAGAGATTATCAA 6112  
 Db 5313 GTCTGACGCTCAGTGAAACGAAACCTCAGTTAAGGATTTTGGTCAAGAGATTATCAA 5372  
 QY 6113 AAGGATCTTACCTAGATCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAA 6172  
 Db 5373 AAGGATCTTACCTAGATCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAA 5432  
 QY 6173 ATATAGTAAACCTTGGTCTGACAGTTACCAATGCTTAAATCTAGTGGGACCTTATCTCAGC 6232  
 Db 5433 ATATAGTAAACCTTGGTCTGACAGTTACCAATGCTTAAATCTAGTGGGACCTTATCTCAGC 5492  
 QY 6233 GATCTGCTATTTCTGTTTCAATCAATGCTTAAATCTAGTGGGACCTTATCTCAGC 6292  
 Db 5493 GATCTGCTATTTCTGTTTCAATCAATGCTTAAATCTAGTGGGACCTTATCTCAGC 5552  
 QY 6293 ACGGAGGGGCTTACCAATCTGGGCTTCAATGCTGCAATGATACCGCGAGACCCAGCTCACC 6352

5553 ACGGAGGGCTTACCATCTGGCCCGACGCTGCTCAATGATACCGCGAGACCCACGCTCACC 5612  
 QY 6353 GGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGAGGAGGCGGAGCGCAGAAAGTGGTCC 6412  
 Db 5613 GGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGAGGAGGCGGAGCGCAGAAAGTGGTCC 5672  
 QY 6413 TCAAACTTTATCCGCTCCATCCAGTCTATTAATTTGTTCCCGGGAAGCTAGAGTAAGTAG 6472  
 Db 5673 TCAAACTTTATCCGCTCCATCCAGTCTATTAATTTGTTCCCGGGAAGCTAGAGTAAGTAG 5732  
 QY 6473 TTAGCGAGTTAATAGTTTGCAGCAAGTTGTTGCTTCCAGGCTACAGGC - TCGTGGTGTACG 6531  
 Db 5733 TTAGCGAGTTAATAGTTTGCAGCAAGTTGTTGCTTCCAGGCTACAGGCATCGTGGTGTACG 5792  
 QY 6532 CTCGTCGTTGGTATGGCTTCAATTCAGCTCCGTTCCCAACGATCAAGCGAGTTACATG 6591  
 Db 5793 CTCGTCGTTGGTATGGCTTCAATTCAGCTCCGTTCCCAACGATCAAGCGAGTTACATG 5852  
 QY 6592 ATCCCCCATGTTGTGCAAAAAAGCGTTAGCTCTTTCGCTCCGATCCGATCGTTGTGTCAGAA 6651  
 Db 5853 ATCCCCCATGTTGTGCAAAAAAGCGTTAGCTCTTTCGCTCCGATCCGATCGTTGTGTCAGAA 5912  
 QY 6652 TAAAGTGGCGCAGTGTATCACTCATGTTAAGGAGCACTGCAATAATCTTCTTACTGT 6711  
 Db 5913 TAAAGTGGCGCAGTGTATCACTCATGTTAAGGAGCACTGCAATAATCTTCTTACTGT 5972  
 QY 6712 CATGCCATCCGTAAGATGCTTTCTGACCTGGTGGTACTCAACCAAGTCAATCTGAGA 6771  
 Db 5973 CATGCCATCCGTAAGATGCTTTCTGACCTGGTGGTACTCAACCAAGTCAATCTGAGA 6032  
 QY 6772 ATAGTGTATGCGCGACCGAGTTGCTCTTTCGCGCGCTCAATACCGGATTAATACCGCGCC 6831  
 Db 6033 ATAGTGTATGCGCGACCGAGTTGCTCTTTCGCGCGCTCAACCAAGTAAATACCGCGCC 6092  
 QY 6832 ACATAGCAGAACTTTTAAAGTGTCTCATATGGAAGAGTCTTTCGCGCGGAGAACTCTC 6891  
 Db 6093 ACATAGCAGAACTTTTAAAGTGTCTCATATGGAAGAGTCTTTCGCGCGGAGAACTCTC 6152  
 QY 6892 AAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCACTGATC 6951  
 Db 6153 AAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCACTGATC 6212  
 QY 6952 TTACGATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAACAGGAAGGCAAAATGC 7011  
 Db 6213 TTACGATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAACAGGAAGGCAAAATGC 6272  
 QY 7012 CGCAAAAAAGGGAATAAGGGCGACACGGAATAATGTTGAATACTATACTTCTCTTTTCA 7071  
 Db 6273 CGCAAAAAAGGGAATAAGGGCGACACGGAATAATGTTGAATACTATACTTCTCTTTTCA 6332  
 QY 7072 ATATTTATGAGCATTTTATCAGGGTTATGTTCTCATGAGCGGATACATATTTGAATGAT 7131  
 Db 6333 ATATTTATGAGCATTTTATCAGGGTTATGTTCTCATGAGCGGATACATATTTGAATGAT 6392  
 QY 7132 TTAGAAAAATAAACAATAAGGGTTCCGCGACATTTCCCGAAAAAGTGCCACTGACGT 7191  
 Db 6393 TTAGAAAAATAAACAATAAGGGTTCCGCGACATTTCCCGAAAAAGTGCCACTGACGT 6452  
 QY 7192 CTAAGAAACCATTTATTCATGACATTAACCTTATAAAAATAGGCGTATCACGAGGCCCTT 7251  
 Db 6453 CTAAGAAACCATTTATTCATGACATTAACCTTATAAAAATAGGCGTATCACGAGGCCCTT 6512  
 QY 7252 TCGTCT 7257  
 Db 6513 TCGTCT 6518

RESULT 15  
 SYNMOVI  
 LOCUS  
 DEFINITION  
 ACCESSION

SYNMOVI 6374 bp DNA linear SYN 27-APR-1993  
 Moloney murine leukemia virus retroviral vector pUXSHD, complete  
 sequence.  
 M64753



```

VERSION      M64753.1 GI:208855
KEYWORDS     cloning vector; hisicidinol dehydrogenase; retroviral vector.
SOURCE       unidentified cloning vector
ORGANISM     other sequences; artificial sequences; vectors.
REFERENCE    1 (bases 1 to 6374)
AUTHORS      Stockschlaeder, M.A., Storb, R., Osborne, W.R. and Miller, A.D.
TITLE        L-hisicidinol provides effective selection of
              retrovirus-vector-transduced keratinocytes without impairing their
              proliferative potential
JOURNAL      Hum. Gene Ther. 2 (1), 33-39 (1991)
MEDLINE      91322161
PUBMED       1650586
COMMENT      Original source text: Synthetic, DNA.
FEATURES     location/Qualifiers
               1..6374
               /organism="unidentified cloning vector"
               /mol_type="genomic DNA"
               /db_xref="taxon:45196"
               175..763
               /note="Mo-MuSV 5' long terminal repeat"
               833..1642
               /note="extended packaging signal"
               1223..1225
               /note="gag ATG start codon to TAG stop codon"
               1655..1994
               /note="simian virus 40 early promoter"
               2004..3308
               /EC_number="1.1.1.23"
               /codon_start=1
               /transl_table=11
               /product="histidinol dehydrogenase"
               /protein_id="AAA73023.1"
               /db_xref="GI:208856"
               /translation="MSFNLLIDWNSCSPEORALLTRPAISASISITRTVSDILDNVK
               TRGDALREYSAKDEKTEVTRALVTPPEIHAAGARLDELQKQMTAAVKNIETPHSAQ
               TLPVDVETQGVRCQVTRVSVGLYIPGGSAPLFSLVLMATPARIAGCQKVLIC
               SPPVADIELYAAQLCGVQEIFNVGGAQIAAALAFSGSEVPKVDKI FPGNFAFVTEAK
               RQVORLDGAAIDPAGSEVLVIADSGATPDFVASDILLSOAEHGPDPVLLTPDAD
               IARKVAVEROLAEPRADTAROLASRLIVTKDLAOCVAISNOYGPBHLIIQTRN
               ARDLVDATISNGSVPLGDWSPESAGDYASGNHVLPTGYTATCSSLGLADFPQRMV
               QELSKAFSAUASITETLUAARUTAHKNVATLRNALKEQA"
               3420..4013
               /note="Mo-MuLV 3' long terminal repeat"

LTR
ORIGIN
Query Match      40.4%; Score 3094.8; DB 12; Length 6374;
Best Local Similarity 71.4%; Pred. No. 0;
Matches 4802; Conservative 0; Mismatches 1382; Indels 538; Gaps 34;

Qy 549 TTTTGAAGAGCCACACCTGAGTTTGGCAAGCTAGCTTAAGTAAACGCCATTTTGGCAAGG 608
Db 174 TTTTGAAGAGCCACACCGCTAGG--TGGCAAGCTAGCTTAAGTAAACGCCATTTTGGCAAGG 231

Qy 609 CATGG-AAAATACATACTGAGATAGAGAGTTTCAGATCAAGTTAGGACACAGAGAC 667
Db 232 CATGGAAAAATACATACTGAGATAGAGAGTTTCAGATCAAGTTAGGACACAGAGAAC 291

Qy 668 AGCAGATATGGGCCAAACAGAGATATCTGTGTAGCAGTTCTCGCCCC--GCTCAGGGCC 726
Db 292 AGCTGAATA----CCAAACAGGATATCTGTGTAGCAGTTCTCGCCCCGCTCAGGGCC 347

Qy 727 AAGAACAGTTGGACAGGAGAT-ATGGGCCAAACAGGATATCTGTGTAGCAGTTCTCCT 785
Db 348 AAGAACAGATGAGACAGCTGAGTGTGGGCCAAACAGGATATCTGTGTAGCAGTTCTCCT 407

Qy 786 GCCCGGCTCAGGGCCAGAGACAGATGTCCTCCAGATGCGGTCCCGCCCTCAGCAGTTTC 845
Db 408 GCCCGGCTCAGGGCCAGAGACAGATGTCCTCCAGATGCGGTCCAGCCCTCAGCAGTTTC 467

Qy 846 TAGAGAACCATCAGATGTTTCCAGGTGCCCCAAGGACCTG-AAATGACCTGTGCGCTTA 904
Db 468 TAGTGAATCATCAGATGTTTCCAGGGTGGCCCCAAGGACCTGAAAATGACCTGTACCTTA 527

```

```

Qy 905 TTTGAACATAACCAATCAGTTTCGCTTCTCGCTTCTGTTTCGCGCGCTTCTGCTCCCGAGCT 964
Db 528 TTTGAACATAACCAATCAGTTTCGCTTCTCGCTTCTGTTTCGCGCGCTTCTGCTCCCGAGCT 587

Qy 965 CAATAAAAGAGCCACAAACCCCTCACTCGGCGCCAGTCTCCGATAGACTGGGTGCGCC 1024
Db 588 CAATAAAAGAGCCACAAACCCCTCACTCGGCGCCAGTCTTCCGATAGACTGGGTGCGCC 647

Qy 1025 CGGTACCCGCTATTCCCAATAAAAGCCCTTCTGCTGTTTGCATCCGAATCGTGTGCTG 1084
Db 648 CGGTACCCGCTATTCCCAATAAAAGCCCTTCTGCTGTTTGCATCCGAATCGTGTGCTGCTG 707

Qy 1085 ATCTTGGAGAGGTCCTCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1144
Db 708 TTTCTTGGAGAGGTCCTCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 767

Qy 1145 GGTTCACACGAGATTTGAGAGACCCCTCCAGGAGACACACGACCCCGCCGCGGAGGTA 1204
Db 768 GGCTCGTCCGGGATTTGGAGACCCCTCCAGGAGACACACCGA-CCACACCGGAGGTA 826

Qy 1205 AGCTGGCCAGCACTTATCTGTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1264
Db 827 AGCTGGCCAGCACTTATCTGTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 886

Qy 1265 CGCTGCTCGGTACTAGTTAGTAACTAGCTCTGTATCTGCGGAGACCCGTGGTGGAACT 1324
Db 887 CGCTGCTCGTACTAGTTAGTAACTAGCTCTGTATCTGCGGAGACCCGTGGTGGAACT 946

Qy 1325 GACGAGTTCCGAGACACCCGCGCGCAACCCCTGGGAGAGTCCACAGGACTTCGCGGGCGCT 1384
Db 947 GACGAGTTCTGAACACACCCGCGCGCAACCCCTGGGAGAGTCCACAGGACTTCGCGGGCGCT 1006

Qy 1385 TTTTGTGGCCGAGCTGAGTCTTAAATCCGATCGTGTGTTAGGACTCTTTTGGTGACACCCC 1444
Db 1007 TTTTGTGGCCGAGCTGAGGAGGAGTGGATGGAATCGGACCC----- 1052

Qy 1445 CTTAGAGAGGAGATATGTTGTTCTGTGAGGAGACGAGAACCTAAACAGTTCCCGCTCC 1504
Db 1053 -----CGTCAGGATATGTTGTTCTGTGAGGAGACGAGAACCTAAACAGTTCCCGCTCC 1107

Qy 1505 GTCTGAATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1561
Db 1108 GTCTGAATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1167

Qy 1562 CAGCATCGTTCTGTTGTTGTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1621
Db 1168 CAGCATCGTTCTGTTGTTGTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1227

Qy 1622 CCCGGGCTAGACTGTTACCACTCCCTTTAAGTTTGAACCTTAGTGTACCTGAGGAAAGATGTCGA 1681
Db 1228 CC-----AGACTGTTACCACTCCCTTTAAGTTTGAACCTTAGTGTACCTGAGGAAAGATGTCGA 1281

Qy 1682 GCGATCGCTCAACACAGTCGGTAGATGTCAAGAGAGAGCTTGGGTACTTCTGCTC 1741
Db 1282 GCGATCGCTCAACACAGTCGGTAGATGTCAAGAGAGAGCTTGGGTACTTCTGCTC 1341

Qy 1742 TGCAGATGCGCAACCTTTAAACGTCGGATGCGCGGAGACGCGACCTTTTAAACCGAGACT 1801
Db 1342 TGCAGATGCGCAACCTTTAAACGTCGGATGCGCGGAGACGCGACCTTTTAAACCGAGACT 1401

Qy 1802 CATCACCCAGGTTAAGATCAAGGTCTTTTCACTTGGCGCGCATGGACACCCAGACCCAGCT 1861
Db 1402 CATCACCCAGGTTAAGATCAAGGTCTTTTCACTTGGCGCGCATGGACACCCAGACCCAGCT 1461

Qy 1862 CCCCTACATCGTACCTGGGAAGCCTTGGCTTTTGAACCCCTCCCTGGGTCAAGCCCTT 1921
Db 1462 CCCCTACATCGTACCTGGGAAGCCTTGGCTTTTGAACCCCTCCCTGGGTCAAGCCCTT 1521

Qy 1922 TGTACACCTTAAGCCTCGCTCTCTTCTTCTTCCATCGCGCCCTCTCTCCCTTGAACC 1981
Db 1522 TGTACACCTTAAGCCTCGCTCTCTTCTTCTTCCATCGCGCCCTCTCTCTCCCTTGAACC 1581

```

1982 TCCGCTTCGACCCCGCCTCGATCCTCCCTTTATCCAGCCCTCCTCCTCTCTAGCGC 2041  
Db |||||  
1582 TCCGCTTCGACCCCGCCTCGATCCTCCCTTTATCCAGCCCTCCTCCTCTAGCGC 1641  
Qy |||||  
2042 CCCCATATGCCCATATGAGATCTTATGCGGACACCCCGCCTTGTGTAATCTCCCTGA 2101  
Db |||||  
1642 CGGAATTCGTAACTCGAGGATCCCGCTGTGGAATGTGTGTCAGTTAGGAGTGTGAAAGT 1701  
Qy |||||  
2102 CCTGTGACATGACAAGAGTTACTAAACAGCCCTCTCTCCAAAGCTCACTTACAGGCTTCTAC 2161  
Db |||||  
1702 CCCAGGCTCCCGAG-----CAGGCAGAAAGTATGAAAGATGCACTCA 1746  
Qy |||||  
2162 TTAGTCCAGCAGAAAGTCTGAGACCTCTGGGGCAGCCTACCAAGAACAACTGACCGCA 2221  
Db |||||  
1747 ATTAGTCAGAACACAGGTGTGAAAGTCCCGAGGCTCCCGCAGCAGCAGAGATATGCARA 1806  
Qy |||||  
2222 CCGGTGGTACCTACCCCTTACCGAGTCCGGGACACAGTGTGGTTCGCCGACCAAGACT 2281  
Db |||||  
1807 GCATGCATCTCAATTAGTCAGCAACCAATAGTCCCGCCCTAACTCCGCCCATCCC----- 1861  
Qy |||||  
2282 AAGAACCTAGAACCTCGCTGGAAGGACCTTACACAGTCTCTGCTGACACCCCGCCCGC 2341  
Db |||||  
1862 --GCCCTTAACCTCCGCCAGTTCGCCCAATCTCCGCCCACTGCTGACATAATTTTTTT 1919  
Qy |||||  
2342 CTCAAGATAGCGGATCGCAGCTTGGATACACCGCCCGCCAGCTGAAGCTGCCGACCCC 2401  
Db |||||  
1920 ATTATGACAGCCGAGCGCGCT-----CGGCTCTGAGCTATTCAGAGTA 1969  
Qy |||||  
2402 GGGGTGGACCATCCTCTCTAGATGCGCATGGGATGGAGCTGTATCATCTCTCTTGGTAG 2461  
Db |||||  
1970 GTGAGGAGGCTTTTGGAGGAGATCCCGGACCATGAGCTTCAATACCTGATGACTG 2029  
Qy |||||  
2462 CACAGCTACAGGTGCTCACTCCGACATCCAGTACCCAGCCAGAGCCCAAGCAGCTGAGC 2521  
Db |||||  
2030 GAACAGCTAGCCCTGAA-----CAGCAGCTGCGCTGCTGACGCTCCGGCA 2079  
Qy |||||  
2522 CCAGCTGGGTGACAGAGTGACCATCACCCTGTAAGGCCAGTCAGGATGGGTACTTCTG 2581  
Db |||||  
2080 TTTCCGCTCTGACAGTATTAACCGGACGGTCAGCGATATCTGGATATGTAAACGC 2139  
Qy |||||  
2582 TAGCTTGTGACAGCAGAGCCAGGTAAGCTCCAAAGCTGCTGATCTACTGGACATCCA 2641  
Db |||||  
2140 GCGGTGACGATGCTGCTGCTGAATACAGCGCTAAATTTGATAAAACAGAAAGTGACAGCG 2199  
Qy |||||  
2642 CCGGCACACTGGTGTGCAACAGATTCAGCGGTAGCGGTAGCGGTACCGACTTCACCT 2701  
Db |||||  
2200 TAGCGTCCCGCTGAAGAGATCGCGCGCGCGCGCGCTGAGCGACGAATTA----- 2256  
Qy |||||  
2702 TCACCATCAGCAGCTCCAGCCAGAGGACATCGCACCTACTACTGCCAGCAATATAGCC 2761  
Db |||||  
2257 --AACAGCGATGACCGCTGCCGTCAAATAATTGAAAGTTCCATTCCGCGCAGAGCT 2314  
Qy |||||  
2762 TCTATCGGTGTTGCGCCCAAGGACCAAGGTG- GAAATCAAAACGAGGTGGCTCAGGATCG 2820  
Db |||||  
2315 ACGCGCTGTAGATGTGGAACCCAGCAGCGGTGCGTTGCCAGCAGTTAAGCGTCCCGT 2374  
Qy |||||  
2821 GGTGATCCGGCTCTGGTGGCTCAGGATCGAGGTCCAACTGTTGAGAGCGGTGAGGT 2880  
Db |||||  
2375 CTCGTCTGTGCTGTATATTCCCGCGGCTCGGCTCCGCTCTTCTCAACGGTGTGAT 2434  
Qy |||||  
2881 GTTGTCAACCTGGCGGCTCCCTGCGCGTGTCTGTCTCGCATCTGGCTTCGATTTCAAC 2940  
Db |||||  
2435 GCTGGCGACCGCGCGCATTTGCGGGATGCCAGAGGTGTTCTGTCTCCCGCGCC 2494  
Qy |||||  
2941 ACATATTTGATGATTTGGTGGAGACAGGACCTTGGAAAAGGTCTTGGTGGATTTGGAGAA 3000  
Db |||||  
2495 CATCGCTGATGAAATCTCTATGCGCGCAAC-----TGTTGGGTGCGAGGA 2542  
Qy |||||  
3001 ATTCACTCAGATAGCAGTACGATTAATCTATGCGCGTCTCTAAAGATAGATTACATA 3060  
Db |||||  
2543 AATCTTTTAACTGCGCGCGCGCAGCGGATTCGCGCTCT-----GGCC 2585  
Qy |||||  
3061 TCGCGAGACAAACGCCAAGAACACATTGTTCTCTGCAATGGACAGCCTGAGACCCGAGAC 3120

2586 TTCCGACGAGCTCCGTACCGAAAGTGATAAATTTTGGCCCGGCAACCGCTTTGTA 2645  
Qy |||||  
3121 ACCGGGCTTATTTTGTGCAAGCCTTACTTCCGCTTCCCCTGTGTGCTTATTTGGGCG 3180  
Db |||||  
2646 ACCGAAGCCAAACGTCAGGTCCAGCAGCGTCTCGACGGCGCGCTATCGAT----- 2696  
Qy |||||  
3181 CAAGGGACCCCGGTCAACGCTCTCCAGTCTTAAGCCCAACACAGCAGCCAGCGCGGACCA 3240  
Db |||||  
2697 -ATGCCAGCGCGGCGCTCTGAAGTACTGGTATCGCA-GACAGCGCGCAACACCGGAT 2754  
Qy |||||  
3241 CCAACAACCGCGCCCAACCATCGCTCGAGCCCTGTCTCTCGCGCCCAAGCGGCTCGG 3300  
Db |||||  
2755 TCGTCGCTTCTGACCTGTCTCTCCAGGCTGAGCACGGCCGGATTTCCAGGTGATCCTGC 2814  
Qy |||||  
3301 CCAGCGCGGGGGGGCGAGTGACACAGAGGGGGCTGGACTTCGCCCTGGATCCCAACTC 3360  
Db |||||  
2815 TGACGCTGATGCTGACATTCGCGCAAGGTGGCGG-----AGGCGGTAGAACGTCACCTG 2870  
Qy |||||  
3361 TGCTACCTGCTGGATGGAATCCTCTTATCTATGTGTCTTCTCACTGCTTGTTCCTG 3420  
Db |||||  
2871 GCGGAACCTGCCCGCGCGGACACCGCCCGGCGGCTGAGCGCAGTCTGCTGATGTG 2930  
Qy |||||  
3421 AGAGTGAAGTTTCAGCAGGAGCGCAGAGCCCGCGGTACACGAGCGCCAGAACACGCTC 3480  
Db |||||  
2931 ACCAAAGATTTAGCGCAGTGGCTG- --CCATCTCTAATCAGTATGGCGCGCAACTTA 2987  
Qy |||||  
3481 TATAAGAGCTCAATCTAGGACGAGAGAGAGTAGGATGTTTTGGACAGAGAGCTGGC 3540  
Db |||||  
2988 ATCATCCAGACGCGCAATTCGCGCGATTTGGTGGATGCGATTTACAGCGCAGGCTCGTA 3047  
Qy |||||  
3541 CGGGACCTGTAGATGGGGGAAAGCCGAGAGGAAGAACCTTCAGAGAGGCTGTACAA 3600  
Db |||||  
3048 TTTCTCGGAGCTGTGTCGCGGATTCGCGGTGATTTAGCTTCCGGAACCAACCATGTT 3107  
Qy |||||  
3601 GAACTGCAGAAAGATAAGATGCGGAGGCTCACTAGTGAGATTGGGATGAAAGCGAGCGC 3660  
Db |||||  
3108 TTACCGACTATGGCTATCTGCTACCTGTTCCAGCTTGGGTAGCGGATTTCCAGAAA 3167  
Qy |||||  
3661 CGGAGGGGAGGGGACAGTGGCTTTACCGAGGTCTCAGTACAGCCACCAAGGACCC 3720  
Db |||||  
3168 CGGATGACCGTTTCAGGA- ACTGTCGAAAGCGGGCTTTTCCGCTTGGCATCAACATTGA 3226  
Qy |||||  
3721 TAGCAGCGCTTCACATGACGAGCCCTGCGCCCTGCTAACTCGACGCGCGCGGATCCG 3780  
Db |||||  
3227 AACATTGGCGCGCGAGAACGTCGACCGCCCTAATAAATGCCGTGA CCGTCCGCGTAAA 3286  
Qy |||||  
3781 GATTAGTCCAATTTTAAAGACGAGATATCAGTGGTCCAGGCTCTAGTTTGACTCAAC 3840  
Db |||||  
3287 CGCCCTCAAGGAGCAAGCATGAGCACTGAAACACTCTCAGCGTCTGCTGACTTAGCCCG 3346  
Qy |||||  
3841 AATATCAACAGCTGAAGCTTATAGATGAGGACCATAGATAAATAAAGATTTTATTTA 3900  
Db |||||  
3347 GATCTGCC-----TAGGCTTTTGCAAAAGCTTATCGATAAATAAAGATTTTATTTA 3400  
Qy |||||  
3901 GTCTCCAGAAAAGGGGGGAAATGAAGACCCCACTGTAGGTTTGGCAGCTAGCTTAAG 3960  
Db |||||  
3401 GTCTCCAGAAAAGGGGGGAAATGAAGACCCCACTGTAGGTTTGGCAGCTAGCTTAAG 3460  
Qy |||||  
3961 TAAACCCCAATTTTCAAGCATGG-AAAATAACAATACTAGATAAGAGTTTCAGATCAA 4019  
Db |||||  
3461 TAAACCCCAATTTTCAAGCATGGAAAATAACAATACTAGATAAGAGTTTCAGATCAA 3520  
Qy |||||  
4020 GGTTAGGAACAGA-GAGACAGCAATATGGGCCAAACAGGATATCTGTGGTAAGAGTT 4078  
Db |||||  
3521 GGTAGGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGGTAAGAGTT 3580  
Qy |||||  
4079 CCGTCCCGC-GCTCAGGGCCAAAGACAGTTGGAACAGGAAATATGGGCCAAACAGGATAT 4137  
Db |||||  
3581 CCGTCCCGGCTCAGGGCCAAAGACAGATGGAACAGCTGAATATGGGCCAAACAGGATAT 3640  
Qy |||||  
4138 CTGTGGTAAGCAGTTCTCTGCCCGCTCAGGGCCAAAGAACAGATGTTCCCGAGATGCGGT 4197



QY 6357 CCAGATTTATCAGCAATAAACCAGCCAGCGGAGCGGCGAGCGAGAGAGTGGTCTGCA 6416  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 5469 CCAGATTTATCAGCAATAAACCAGCCAGCGGAGCGGCGAGCGAGAGTGGTCTGCA 5528  
 QY 6417 ACTTTATCCGCTCCATCAGTCTATTAATTTGTCGGGGAAGCTAGATAAGTAGTTCG 6476  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 5529 ACTTTATCCGCTCCATCAGTCTATTAATTTGTCGGGGAAGCTAGATAAGTAGTTCG 5588  
 QY 6477 CCAGTTAATAGTTTGGCGCAACGTTGTCATTTGCTACAGGC-TGTTGGTGTACGCTCG 6535  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 5589 CCAGTTAATAGTTTGGCGCAACGTTGTCATTTGCTACAGGCATCGTGTGTACGCTCG 5648  
 QY 6536 TCGTTTGGTATGGCTTTCATTCAGCTCCGGTTCCTCAACGATCAAGCGAGTTACATGATCC 6595  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 5649 TCGTTTGGTATGGCTTTCATTCAGCTCCGGTTCCTCAACGATCAAGCGAGTTACATGATCC 5708  
 QY 6596 CCATGTTGTGCAAAAAGGGTTAGTCTCTTCGGTCTCCGATCGTTGTCTAGAGTAAG 6655  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 5709 CCATGTTGTGCAAAAAGGGTTAGTCTCTTCGGTCTCCGATCGTTGTCTAGAGTAAG 5768  
 QY 6656 TTGGCCGCAAGTTATCACTCATGTTATGGCAGCACTGCATAATTTCTTACTGTCTAG 6715  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 5769 TTGGCCGCAAGTTATCACTCATGTTATGGCAGCACTGCATAATTTCTTACTGTCTAG 5828  
 QY 6716 CCATCCGTAAAGATGCTTTTCTGTGACTGGTGTGAGTACTCAACCAAGTCAATTCAGAGTAAG 6775  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 5829 CCATCCGTAAAGATGCTTTTCTGTGACTGGTGTGAGTACTCAACCAAGTCAATTCAGAGTAAG 5888  
 QY 6776 TGTATCGGCGACCGAGTTGCTCTTCGCCGGGTCAATACGGGATTAATACCGGCCACAT 6835  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 5889 TGTATCGGCGACCGAGTTGCTCTTCGCCGGGTCAATACGGGATTAATACCGGCCACAT 5948  
 QY 6836 AGCAGAACTTTAAAGTCTCATCTTGGAAACGTTCTTCGGGGGAAACCTCTCAAG 6895  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 5949 AGCAGAACTTTAAAGTCTCATCTTGGAAACGTTCTTCGGGGGAAACCTCTCAAG 6008  
 QY 6896 ATCTTACCCTGTTGAGATCCAGTTCGATGTAAACCACTCGTGCACCCCACTCATCTTCA 6955  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 6009 ATCTTACCCTGTTGAGATCCAGTTCGATGTAAACCACTCGTGCACCCCACTCATCTTCA 6068  
 QY 6956 GCATCTTTTACTTTTCAACGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGA 7015  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 6069 GCATCTTTTACTTTTCAACGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGA 6128  
 QY 7016 AAAAGGGAATAGGCGGACAGGAATGTTGAAATCTCATCTCTTCTCTTTTCAATAT 7075  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 6129 AAAAGGGAATAGGCGGACAGGAATGTTGAAATCTCATCTCTTCTCTTTTCAATAT 6188  
 QY 7076 TATTGAAGCAATTTATCAGGGTTATTGTTCTCATGCGGATACATATTTGAATGTAATTTAG 7135  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 6189 TATTGAAGCAATTTATCAGGGTTATTGTTCTCATGCGGATACATATTTGAATGTAATTTAG 6248  
 QY 7136 AAAATAAACAATAGGGTTCGCGCAATTTCCCGGAAAGTGCCACCTGACGCTAA 7195  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 6249 AAAATAAACAATAGGGTTCGCGCAATTTCCCGGAAAGTGCCACCTGACGCTAA 6308  
 QY 7196 GAAACCATTTATCATGATTAACCTATATAAGGGTATACAGAGGCCCTTTCTGT 7255  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 6309 GAAACCATTTATCATGATTAACCTATATAAGGGTATACAGAGGCCCTTTCTGT 6368  
 QY 7256 CT 7257  
 Db ||  
 6369 CT 6370

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: May 25, 2005, 05:30:20 ; Search time 2422 Seconds  
(without alignments)  
18707.556 Million cell updates/sec

Title: US-10-006-771B-1  
Perfect score: 7654  
Sequence: 1 aagcttcagctgcaggt.....gtaaacagcggccagtgcc 7654

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04.\*  
1: Geneseqn1980s.\*  
2: Geneseqn1990s.\*  
3: Geneseqn2000s.\*  
4: Geneseqn2001as.\*  
5: Geneseqn2001bs.\*  
6: Geneseqn2002as.\*  
7: Geneseqn2002bs.\*  
8: Geneseqn2003as.\*  
9: Geneseqn2003bs.\*  
10: Geneseqn2003cs.\*  
11: Geneseqn2003ds.\*  
12: Geneseqn2004as.\*  
13: Geneseqn2004bs.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7654	100.0	7654	8: ABX15565	Abx15565 Retrovira
2	7654	100.0	7654	10: ABX13168	Abx13168 Retrovira
3	3349	43.8	9320	6: ABS56664	Abx56664 Plasmid p
4	3307.8	43.2	6644	2: AAX33181	Aax33181 Base sequ
5	3244.4	42.4	6505	9: ADA12886	Ada12886 Murine MS
6	3223.6	42.1	5903	13: ADQ80674	Adq80674 Bcl-XL ex
7	3213	42.0	5782	6: ABR85887	ABR85887 DNA sequ
8	3183.2	41.6	7372	2: AAX33182	Aax33182 Base sequ
9	3139.6	41.0	6253	9: AAL57215	Aal57215 MIG retro
10	3130.2	40.9	7257	13: ADQ80673	Adq80673 Bcl-XL ex
11	3111	40.6	5364	2: AAT13393	Aat13393 Hybrid ve
12	3103.4	40.5	6444	6: AAI67595	Aai67595 Nucleotid
13	3096	40.4	6522	2: AAX90482	Aax90482 Plasmid r
14	3093	40.4	6221	3: AAX34935	Aax34935 Retrovira
15	3093	40.4	6221	4: AAF30944	Aaf30944 Vector us
16	3092.4	40.4	6365	2: AAO41173	Aao41173 Plasmid L
17	3083.4	40.3	5323	2: AAT13390	Aat13390 Hybrid ve
18	3080.2	40.2	5294	2: AAT13391	Aat13391 Hybrid ve
19	3077	40.2	5292	2: AAT13392	Aat13392 Hybrid ve
20	3062.2	40.0	6145	2: AAV83182	Aav83182 Intermedi

21	3062.2	40.0	6145	5	AAF85611	Aaf85611 M-MuLV-ba	
22	3050.2	39.9	6620	6	ADE82653	Ade82653 Plasmid v	
23	3006.8	39.3	6620	9	ADA12887	Ada12887 Murine le	
24	2899.2	37.9	6141	2	AAX90481	Aax90481 Plasmid r	
25	2814	36.8	6046	12	ADG83262	Adg83262 Clone pLX	
26	2783.8	36.4	7797	2	AAX33180	Aax33180 Cowpox vi	
27	2687.4	35.1	5874	2	AAX90484	Aax90484 Plasmid r	
28	2683.4	35.1	5865	2	AAV04002	Aav04002 Retrovira	
29	2675.4	35.0	6283	4	AAF83147	Aaf83147 Complete	
30	2672.4	34.9	7086	6	ABK12523	Abk12523 Tetracycl	
31	2639	34.5	7165	2	AAX90483	Aax90483 Plasmid r	
32	2634.8	34.4	5421	3	AAC68299	Aac68299 SV40/APPA	
33	2634.8	34.4	6116	3	AAC68297	Aac68297 R15/APPA	
34	2634.8	34.4	6708	3	AAC68295	Aac68295 R15/APPA	
35	2634	34.4	6143	9	ADA12885	Ada12885 Murine re	
36	2633.2	34.4	7560	4	AAC85599	Aac85599 Plasmid p	
C	37	2633.2	34.4	7560	9	ADA09849	Ada09849 piggyBac
38	2630.4	34.4	5715	3	AAZ34936	Aaz34936 Retroviri	
39	2630.4	34.4	5715	4	AAF30945	Aaf30945 Vector us	
40	2627.6	34.3	4163	2	AAQ32349	Aaq32349 Template	
41	2624.4	34.3	5465	2	AAZ20088	Aaz20088 Plasmid p	
C	42	2624	34.3	3309	6	AAK98929	Aak98929 Human pSG
43	2623.4	34.3	9164	2	AAX82259	Aax82259 Beta-doma	
44	2623.4	34.3	11846	2	AAX82261	Aax82261 Factor VI	
45	2623.4	34.3	12022	2	AAX82260	Aax82260 Factor VI	

## ALIGNMENTS

## RESULT 1

ABX15565  
ID ABX15565 standard; DNA; 7654 BP.

XX ABX15565;

DT 22-APR-2003 (first entry)

DE Retroviral vector expressing a chimaeric hMN14/T-cell receptor.

XX Retroviral vector; ds; T-cell receptor; hMN14; antibody; IgTCR;

KW cytosolic; dermatological; neuroprotective; immunostimulant; GD3;

KW ganglioside antigen; MB3.6; PSMA; tumour; 3D8; 4D4; 3E11; cancer;

KW prostate-specific membrane antigen; zeta signalling chain; human; cancer;

KW melanoma; neuroendocrine tumour; prostate cancer; small cell lung cancer;

KW mouse; C8alpha hinge.

XX Homo sapiens.

OS Mus sp.

OS Retrovirus sp.

OS Synthetic.

OS Chimeric.

Key Location/Qualifiers

CDS 2428..3759

/\*tag= a

/product= "IgTCR"

US2002132983-A1.

19-SEP-2002.

10-DEC-2001; 2001US-00006773.

30-NOV-2000; 2000US-0250087P.

30-NOV-2000; 2000US-0250089P.

(JUNG/) JUNGHANS R P.

Junghans RP;

WPI; 2003-208946/20.

P-ESDB; ABG74240.



XX New chimeric molecule useful in treating patients with disorders, such as  
PT melanoma, neuroendocrine disorders, prostate and small cell lung cancer  
PT comprise GD3 and/or PSMA binding domains of antibody.  
XX  
PS Disclosure; Page 3-7; 35pp; English.

XX  
CC The invention relates to a chimeric molecule comprising the GD3  
CC (ganglioside antigen) binding domain of antibody MB3.6, with any of 3  
CC variable gene sequences, or the PSMA (prostate-specific membrane antigen)  
CC binding domain of antibody 3D8, 4D4 and 3E11, with variable gene  
CC sequences, the zeta signalling chain of the T cell receptor and an  
CC intervening CD8alpha hinge in which cysteine residues have been mutated.  
CC The chimeric molecules expressed in T cells or NK cells or other  
CC effector cells are useful in treating patients with cancers expressing  
CC the GD3 (MB3.6 derivatives) or PSMA antigen (3D8, 4D4, 3E11 derivatives),  
CC and/or together with each other or with heterologous constructs to engage  
CC additional stimulatory and functional properties of the effector cells to  
CC enhance the antitumour therapeutic efficacy (claimed). They are  
CC particularly useful in disorders including melanoma, neuroendocrine  
CC tumours and prostate and small cell lung cancer. The present sequence is  
CC a retroviral vector encoding the hMN14 antibody (specific to CEA antigen)  
CC in a fusion protein with the modified CD8alpha hinge and the T-cell  
CC receptor zeta chain (IgTCR). The hMN14 antibody coding region is replaced  
CC with the MB3.6, 3D8, 4D4 or 3E11 genes of the invention  
XX

SQ Sequence 7654 BP; 1832 A; 2067 C; 1953 G; 1802 T; 0 U; 0 Other;

Query Match 100.0%; Score 7654; DB 8; Length 7654;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 7654; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGCTTGCACTGCTGAGTGCAGTCTAGGCACATTAAGAAAACATACTAACCAAGCT 60  
DB 1 AAGCTTGCACTGCTGAGTGCAGTCTAGGCACATTAAGAAAACATACTAACCAAGCT 60  
QY 61 GCAGCGGACAGTGAAGAAACCGTTAAACGGTTGTTTAAATAAATCAATATATTT 120  
DB 61 GCAGCGGACAGTGAAGAAACCGTTAAACGGTTGTTTAAATAAATCAATATATTT 120  
QY 121 AGAGTCATTTCTTTGGTAGGAAGTACATTTGGCAGTAAAGGAGCCCAAGCAATCTGTG 180  
DB 121 AGAGTCATTTCTTTGGTAGGAAGTACATTTGGCAGTAAAGGAGCCCAAGCAATCTGTG 180  
QY 181 GAAAGCCAGGCTGGAGGCCAGCAGTTTGCATCCCTCTCGCGTGTACCTAAGGGTTT 240  
DB 181 GAAAGCCAGGCTGGAGGCCAGCAGTTTGCATCCCTCTCGCGTGTACCTAAGGGTTT 240  
QY 241 CTTAATTGTGTGTTTCTAAATCTTCCAGAGGTTTGTCTCAATTCACCTTCCACTTCGGTG 300  
DB 241 CTTAATTGTGTGTTTCTAAATCTTCCAGAGGTTTGTCTCAATTCACCTTCCACTTCGGTG 300  
QY 301 CACAATACTTGGACGGGATTTACTGTCTTAGCATCTATCGGTGGCCCTTCGATTGAGGC 360  
DB 301 CACAATACTTGGACGGGATTTACTGTCTTAGCATCTATCGGTGGCCCTTCGATTGAGGC 360  
QY 361 TGAACCTGAGGCCACTTCTTCTAGCTTGTAGGAGAGCACAAGCACAAGAGAGGCTGA 420  
DB 361 TGAACCTGAGGCCACTTCTTCTAGCTTGTAGGAGAGCACAAGCACAAGAGAGGCTGA 420  
QY 421 CCGGGCAGACCTGTGGGCAATTTTAAACAGGGCCCTCTCGGTCTGTGGGAGGAGGCTTA 480  
DB 421 CCGGGCAGACCTGTGGGCAATTTTAAACAGGGCCCTCTCGGTCTGTGGGAGGAGGCTTA 480  
QY 481 CATAAGGTGCAATTAAGAATATAAATAAGCCCATATCAATTTGTCTCTTTTAA 540  
DB 481 CATAAGGTGCAATTAAGAATATAAATAAGCCCATATCAATTTGTCTCTTTTAA 540  
QY 541 AGCTCAAGTTTGAAGACCCCACTGTAGGTTTGGCAAGCTAGCTTAAGTAACGCCATT 600  
DB 541 AGCTCAAGTTTGAAGACCCCACTGTAGGTTTGGCAAGCTAGCTTAAGTAACGCCATT 600  
QY 601 TTGCAAGGCAATGGAATAATACATACTAGCAATAGAGAAGTTTCAGATCAAGGTTAGGAACA 660

DB 601 TTGCAAGGCAATGGAATAATACATACTAGCAATAGAGAAGTTTCAGATCAAGGTTAGGAACA 660  
QY 661 GAGAGACAGCAGCAATATGGGCCAAACAGAGATATCTGTGTAAAGAGTTCTCTCGCCCGCTC 720  
DB 661 GAGAGACAGCAGCAATATGGGCCAAACAGAGATATCTGTGTAAAGAGTTCTCTCGCCCGCTC 720  
QY 721 AGGGCCAAAGACAGTTGGAAACAGGAGATATGGGCCAAACAGAGATATCTGTGTAAAGAG 780  
DB 721 AGGGCCAAAGACAGTTGGAAACAGGAGATATGGGCCAAACAGAGATATCTGTGTAAAGAG 780  
QY 781 TTCTCTGCCCCGCTCAGGCGCCAAAGAACAGATGGTCCCCAGATGCGGTCCCGCCCTCAGCA 840  
DB 781 TTCTCTGCCCCGCTCAGGCGCCAAAGAACAGATGGTCCCCAGATGCGGTCCCGCCCTCAGCA 840  
QY 841 GTTCTTAGAACAACATCAGATGTTTCCAGGGTGCCCAAGGACCTTGAATGACCTGTGTC 900  
DB 841 GTTCTTAGAACAACATCAGATGTTTCCAGGGTGCCCAAGGACCTTGAATGACCTGTGTC 900  
QY 901 CTTATTTGAACTAAACCAATCAGTTTCGCTTCTCGCTTCTGTTCGCGGCTTCTGTCTCCCG 960  
DB 901 CTTATTTGAACTAAACCAATCAGTTTCGCTTCTCGCTTCTGTTCGCGGCTTCTGTCTCCCG 960  
QY 961 AGCTCAATAAAGAGCCCAACACCCCTCACTCGGCGCGCAGTCTCCGATAGACTGCGT 1020  
DB 961 AGCTCAATAAAGAGCCCAACACCCCTCACTCGGCGCGCAGTCTCCGATAGACTGCGT 1020  
QY 1021 GCGCGGGTACCCGTTATTCCTCAATAAAGCCTCTTCTGTGTGATCCGAATCGTGAGCTC 1080  
DB 1021 GCGCGGGTACCCGTTATTCCTCAATAAAGCCTCTTCTGTGTGATCCGAATCGTGAGCTC 1080  
QY 1081 GCTGATCCTTGGGAGGGTCTCTCAGATTGATTGACTGCCACCTCGGGGGTCTTTCATT 1140  
DB 1081 GCTGATCCTTGGGAGGGTCTCTCAGATTGATTGACTGCCACCTCGGGGGTCTTTCATT 1140  
QY 1141 TGGAGGTTTCCACCCGAGATTGGAGACCCCTGCGGAGGACCAACGACCCCGCGCGGA 1200  
DB 1141 TGGAGGTTTCCACCCGAGATTGGAGACCCCTGCGGAGGACCAACGACCCCGCGCGGA 1200  
QY 1201 GGTAAAGTGGCGAGCAACTTATCTGTGTCTGTCGATTTGTCTATGACTGATTT 1260  
DB 1201 GGTAAAGTGGCGAGCAACTTATCTGTGTCTGTCGATTTGTCTATGACTGATTT 1260  
QY 1261 TATGCGCTGCGTGGTACTAGTTAGCTAACTAGCTCTGTATCTCTGGCGGACCCGCTGGTG 1320  
DB 1261 TATGCGCTGCGTGGTACTAGTTAGCTAACTAGCTCTGTATCTCTGGCGGACCCGCTGGTG 1320  
QY 1321 AACTGACAGTTTCCGGAACACCCGCGGCAACCCCTGGGAGACGTCCTCAGGAGCTTCGGGG 1380  
DB 1321 AACTGACAGTTTCCGGAACACCCGCGGCAACCCCTGGGAGACGTCCTCAGGAGCTTCGGGG 1380  
QY 1381 CCGTTTTTGTGGCCGACCTGAGTCTTAAATCCGATCGTTTGGGACTCTTTTGTGTCAC 1440  
DB 1381 CCGTTTTTGTGGCCGACCTGAGTCTTAAATCCGATCGTTTGGGACTCTTTTGTGTCAC 1440  
QY 1441 CCCCCCTTAGAGGAGGATATGTGGTTCTGGTAGGAGACGAGAACCTTAAACAGTTTCCCGC 1500  
DB 1441 CCCCCCTTAGAGGAGGATATGTGGTTCTGGTAGGAGACGAGAACCTTAAACAGTTTCCCGC 1500  
QY 1501 CTCCTGCTGAATTTTGTCTTGGGACCCGAGCCGCGCGCTCTTGTCTGCT 1560  
DB 1501 CTCCTGCTGAATTTTGTCTTGGGACCCGAGCCGCGCGCTCTTGTCTGCT 1560  
QY 1561 GCAGCATCGTTCTGTGTTGCTCTCTGCTGACTGTGTTTCTGTATTTGTCTGAAAATATGG 1620  
DB 1561 GCAGCATCGTTCTGTGTTGCTCTCTGCTGACTGTGTTTCTGTATTTGTCTGAAAATATGG 1620  
QY 1621 GCCCGGGCTAGACTGTGTTACCACTCCCTTAAGTTTGAACCTTAGGTACCTTGAAGATGTCG 1680  
DB 1621 GCCCGGGCTAGACTGTGTTACCACTCCCTTAAGTTTGAACCTTAGGTACCTTGAAGATGTCG 1680  
QY 1681 AGCGGATCGCTCACAAACAGTGGTGTAGATGTCAAGAGAGAGCTTGGGGTACCTTCTGCT 1740

Db 1681 AGCGGATCGCTCACACCAGTCCGTAGATGTCAAGAGAGAGCGTTGGGTTACCTTCTGCT 1740  
Qy 1741 CTGCAGAAATGCCAACCTTTAACTCGGATGCGCGAGAGCGCACCTTTAAACCGAGACC 1800  
Db 1741 CTGCAGAAATGCCAACCTTTAACTCGGATGCGCGAGAGCGCACCTTTAAACCGAGACC 1800  
Qy 1801 TCATCACCCAGGTTAAGATCAAGGTCCTTTTACCTGGSCCGCATGGACACCCAGACGAG 1860  
Db 1801 TCATCACCCAGGTTAAGATCAAGGTCCTTTTACCTGGSCCGCATGGACACCCAGACGAG 1860  
Qy 1861 TCCCTCATACGTGACCTCGGAAGCTTGGCTTTTGAACCCCTCCCTGGGTCAGGCCCT 1920  
Db 1861 TCCCTCATACGTGACCTCGGAAGCTTGGCTTTTGAACCCCTCCCTGGGTCAGGCCCT 1920  
Qy 1921 TTGTACACCCTAAGCTCCGCTCTCTCTCTCCATCGCCCGGCTCTCTCCCTTTGAAC 1980  
Db 1921 TTGTACACCCTAAGCTCCGCTCTCTCTCTCCATCGCCCGGCTCTCTCCCTTTGAAC 1980  
Qy 1981 CTCTCTGTTGACCCCGCTCGATCCTCTCTTTATCGAGCGCTCATCTCTCTTAGGCG 2040  
Db 1981 CTCTCTGTTGACCCCGCTCGATCCTCTCTCTTTATCGAGCGCTCATCTCTCTTAGGCG 2040  
Qy 2041 CCCCATATGSCCATATCAGATCTTATATGGGCAACCCCGCCCTTGTAACTTCCCTG 2100  
Db 2041 CCCCATATGSCCATATCAGATCTTATATGGGCAACCCCGCCCTTGTAACTTCCCTG 2100  
Qy 2101 ACCCTGACATGACAAAGGTTACTTAACAGCCCTCTCTCCAAGCTCACCTTACAGGCTTCTA 2160  
Db 2101 ACCCTGACATGACAAAGGTTACTTAACAGCCCTCTCTCCAAGCTCACCTTACAGGCTTCTA 2160  
Qy 2161 CTTAGTCAGACAAGAGTCTGAGACCTCTGGGCGAGCCTTACCAAGAACTTGAAC 2220  
Db 2161 CTTAGTCAGACAAGAGTCTGAGACCTCTGGGCGAGCCTTACCAAGAACTTGAAC 2220  
Qy 2221 ACCGCTGTACTCTACCTTACCGAGTCGGCGACACAGTGTGGTCCGCCACACCGAC 2280  
Db 2221 ACCGCTGTACTCTACCTTACCGAGTCGGCGACACAGTGTGGTCCGCCACACCGAC 2280  
Qy 2281 TAAGAACTTAGAACCTCTGCTGAAAGGACCTTACACAGTCTCTGTGACACCCCGACCGC 2340  
Db 2281 TAAGAACTTAGAACCTCTGCTGAAAGGACCTTACACAGTCTCTGTGACACCCCGACCGC 2340  
Qy 2341 CCTCAAAGTAGACGGCATCGAGCTTGATACACGCGCCCACTGAGAGGTCGCGACCC 2400  
Db 2341 CCTCAAAGTAGACGGCATCGAGCTTGATACACGCGCCCACTGAGAGGTCGCGACCC 2400  
Qy 2401 CGGGGTTGACCATCTCTAGACTGCCATGGATGGAGCTGTATCATCTCTTCTTGTA 2460  
Db 2401 CGGGGTTGACCATCTCTAGACTGCCATGGATGGAGCTGTATCATCTCTTCTTGTA 2460  
Qy 2461 GCAACAGCTACAGGTGTCCACTCCGACATCCAGCTGACCCAGAGCCCAAGCAGCTTACG 2520  
Db 2461 GCAACAGCTACAGGTGTCCACTCCGACATCCAGCTGACCCAGAGCCCAAGCAGCTTACG 2520  
Qy 2521 GCCAGCTGGGTGACAGAGTCAACATCACTGTAAAGCCAGTCAAGGATGTGGGTACTTCT 2580  
Db 2521 GCCAGCTGGGTGACAGAGTCAACATCACTGTAAAGCCAGTCAAGGATGTGGGTACTTCT 2580  
Qy 2581 GTAGCTTGGTACACGAGAGCCAGGTAAGGCTCCAAAGCTGCTGATCTACTGACATCC 2640  
Db 2581 GTAGCTTGGTACACGAGAGCCAGGTAAGGCTCCAAAGCTGCTGATCTACTGACATCC 2640  
Qy 2641 ACCGGCACACTGGTGTGCCAAGAGATTACGGGTAGCGGTAGCGGTACCGACTTACCC 2700  
Db 2641 ACCGGCACACTGGTGTGCCAAGAGATTACGGGTAGCGGTAGCGGTACCGACTTACCC 2700  
Qy 2701 TTCAACCATCAGACGCTCCAGCCAGAGGACATCGCCACTTACTACTGACCAATATAGC 2760  
Db 2701 TTCAACCATCAGACGCTCCAGCCAGAGGACATCGCCACTTACTACTGACCAATATAGC 2760  
Qy 2761 CTCTATCGGTCTGGCCAAAGGACCAAGGTGAAATCAAAACAGGTGGCTCAGGATCG 2820  
Db 2761 CTCTATCGGTCTGGCCAAAGGACCAAGGTGAAATCAAAACAGGTGGCTCAGGATCG 2820

Qy 2821 GGTGGATCCGCTCTGGTGGCTCAGGATCGAGGTCCAACCTGGTGAAGCGGTGAGGT 2880  
Db 2821 GGTGGATCCGCTCTGGTGGCTCAGGATCGAGGTCCAACCTGGTGAAGCGGTGAGGT 2880  
Qy 2881 GTTGTGCAACTGCGCGGTCCCTGCGCTGCTCTCGCATCTGGCTCGATTTGATTTTACC 2940  
Db 2881 GTTGTGCAACTGCGCGGTCCCTGCGCTGCTCTCGCATCTGGCTCGATTTGATTTTACC 2940  
Qy 2941 ACATATTGGATGAGTTGGGTGAGACAGGACCTCGAAAAGGTCTTGAGTGATTTGAGAA 3000  
Db 2941 ACATATTGGATGAGTTGGGTGAGACAGGACCTCGAAAAGGTCTTGAGTGATTTGAGAA 3000  
Qy 3001 ATTCAATCAGATAGCAGTACGATTAACTATGCGCGCTCTTAAAGGATAGATTACAATA 3060  
Db 3001 ATTCAATCAGATAGCAGTACGATTAACTATGCGCGCTCTTAAAGGATAGATTACAATA 3060  
Qy 3061 TCGCGAGACACGCGCAAGAACACATTTGCTCGAATGGACAGCTGAGACCCGAGAC 3120  
Db 3061 TCGCGAGACACGCGCAAGAACACATTTGCTCGAATGGACAGCTGAGACCCGAGAC 3120  
Qy 3121 ACCGGGTCTATTTTGTGCAAGCTTTACTTTCGGCTTCCCTGCTTGTATTGTTGGGC 3180  
Db 3121 ACCGGGTCTATTTTGTGCAAGCTTTACTTTCGGCTTCCCTGCTTGTATTGTTGGGC 3180  
Qy 3181 CAAGGACCCCGGTCAACCGTCTCCAGTCTTAAGCCCAACACGACGCCAGCCGCGACCA 3240  
Db 3181 CAAGGACCCCGGTCAACCGTCTCCAGTCTTAAGCCCAACACGACGCCAGCCGCGACCA 3240  
Qy 3241 CCAACACGCGCGCCCAACGATCGCTCGAGCCCTGCTCTGCGCCACAGAGCGGCTCGG 3300  
Db 3241 CCAACACGCGCGCCCAACGATCGCTCGAGCCCTGCTCTGCGCCACAGAGCGGCTCGG 3300  
Qy 3301 CCAGCGCGGGGCGCAGTGCACACGAGGGGCTGGACTTTCGCTCGATCCCAAACTC 3360  
Db 3301 CCAGCGCGGGGCGCAGTGCACACGAGGGGCTGGACTTTCGCTCGATCCCAAACTC 3360  
Qy 3361 TGCTACCTGCTGGATGGAATCTCTTCACTATATGGTGTCAATCTCACTGCTTGTTCCTG 3420  
Db 3361 TGCTACCTGCTGGATGGAATCTCTTCACTATATGGTGTCAATCTCACTGCTTGTTCCTG 3420  
Qy 3421 AGAGTGAAGTTTACAGAGGCGAGAGCCCGCCCGTACACAGAGGGCCAGAACGAGCTC 3480  
Db 3421 AGAGTGAAGTTTACAGAGGCGAGAGCCCGCCCGTACACAGAGGGCCAGAACGAGCTC 3480  
Qy 3481 TATAACGAGCTCAATCTAGGACGAAGAGAGAGTACGATGTTTTGGACAAGAGAGCTGCG 3540  
Db 3481 TATAACGAGCTCAATCTAGGACGAAGAGAGAGTACGATGTTTTGGACAAGAGAGCTGCG 3540  
Qy 3541 CGGACCTCTGAGATGGGGGAAAGCCCGAAGAGGAAGAACCTTCAGGAAGGCTGTACAA 3600  
Db 3541 CGGACCTCTGAGATGGGGGAAAGCCCGAAGAGGAAGAACCTTCAGGAAGGCTGTACAA 3600  
Qy 3601 GAATCTCAGAAAGATAGATGGCGGAGGCTTACAGTGAAGTTGGATGAAAGGCGAGCGC 3660  
Db 3601 GAATCTCAGAAAGATAGATGGCGGAGGCTTACAGTGAAGTTGGATGAAAGGCGAGCGC 3660  
Qy 3661 CGGAGGGCAAGGGCGAGATGGCTTTTACAGGGTCTCAGTACAGCCCAAGAGACCC 3720  
Db 3661 CGGAGGGCAAGGGCGAGATGGCTTTTACAGGGTCTCAGTACAGCCCAAGAGACCC 3720  
Qy 3721 TACGACCCCTTCACTGACGGCCCTGCGCTTAACTTCGACCGCGCGCGGATCCG 3780  
Db 3721 TACGACCCCTTCACTGACGGCCCTGCGCTTAACTTCGACCGCGCGCGGATCCG 3780  
Qy 3781 GATTAGTCCAATTTGTTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTGACTCAAC 3840  
Db 3781 GATTAGTCCAATTTGTTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTGACTCAAC 3840  
Qy 3841 AATATCACCGCTGAAGCCTATAGAGTACGAGCCATAGTAAATATAAGATTTTATTTA 3900  
Db 3841 AATATCACCGCTGAAGCCTATAGAGTACGAGCCATAGTAAATATAAGATTTTATTTA 3900



Db 6061 CTCAGTGGAAACGAACTCAGTTAAGGATTTTGGTCATGAGATTATCAAAAAGGATCT 6120  
Qy 6121 TCACCTAGATCCTTTTAAATTAATAAGATTTTAAATCAATCTAAAGTATATATAGT 6180  
Db 6121 TCACCTAGATCCTTTTAAATTAATAAGATTTTAAATCAATCTAAAGTATATATAGT 6180  
Qy 6181 AAATCTGCTGACAGATTACCAATGCTTAATCAGTGGAGCCCTATCTCAGCGATCTGTC 6240  
Db 6181 AAATCTGCTGACAGATTACCAATGCTTAATCAGTGGAGCCCTATCTCAGCGATCTGTC 6240  
Qy 6241 TATTTTGGTTTCATCATAGTTCCTGACTCCCGCTCGTGTAGATAAATACGATACGGGAGG 6300  
Db 6241 TATTTTGGTTTCATCATAGTTCCTGACTCCCGCTCGTGTAGATAAATACGATACGGGAGG 6300  
Qy 6301 CGTTACCATCTGGCCCCAGTGTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAG 6360  
Db 6301 CGTTACCATCTGGCCCCAGTGTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAG 6360  
Qy 6361 ATTTATCAGCAATAAACACGAGCCGGAAGGGCGAGCGAGAGTGGTCTCTGCAACTT 6420  
Db 6361 ATTTATCAGCAATAAACACGAGCCGGAAGGGCGAGCGAGAGTGGTCTCTGCAACTT 6420  
Qy 6421 TATCGGCTCCATCCAGTCTAATTAATTTGTCGGGAAGCTAGAGTAAGTGTTCGCCAG 6480  
Db 6421 TATCGGCTCCATCCAGTCTAATTAATTTGTCGGGAAGCTAGAGTAAGTGTTCGCCAG 6480  
Qy 6481 TTAATAGTTTCGCCAACCGTTGTGCCATTGCTACAGGCTCGTGGTGTACAGCTCGTGGT 6540  
Db 6481 TTAATAGTTTCGCCAACCGTTGTGCCATTGCTACAGGCTCGTGGTGTACAGCTCGTGGT 6540  
Qy 6541 TGGTATGCTTCATTCAGCTCGGTTCCCAACGATCAAGGGAGTTACATGATCCGCCAT 6600  
Db 6541 TGGTATGCTTCATTCAGCTCGGTTCCCAACGATCAAGGGAGTTACATGATCCGCCAT 6600  
Qy 6601 GTTGTGCAAAAGCGGTAGCTCTTCGGTCCCTCGATCGTGTGCAAGTAAGTTGGC 6660  
Db 6601 GTTGTGCAAAAGCGGTAGCTCTTCGGTCCCTCGATCGTGTGCAAGTAAGTTGGC 6660  
Qy 6661 CGCAGTGTATCACTCATGTTATGGCAGCACTGCATAATTTCTTACTGTATGCCATC 6720  
Db 6661 CGCAGTGTATCACTCATGTTATGGCAGCACTGCATAATTTCTTACTGTATGCCATC 6720  
Qy 6721 CGTAAGATGCTTTCTGTGACTGTGAGTACTCAACCAAGTCATTTCTGAGAAATAGTGTAT 6780  
Db 6721 CGTAAGATGCTTTCTGTGACTGTGAGTACTCAACCAAGTCATTTCTGAGAAATAGTGTAT 6780  
Qy 6781 GCGGCGACGAGTTGCTCTTCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAG 6840  
Db 6781 GCGGCGACGAGTTGCTCTTCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAG 6840  
Qy 6841 AACTTTAAAGTGTCTCATCATTTGGAACACGTTCTTCCGGGGGAAACCTCTCAAGATCTT 6900  
Db 6841 AACTTTAAAGTGTCTCATCATTTGGAACACGTTCTTCCGGGGGAAACCTCTCAAGATCTT 6900  
Qy 6901 ACCGCTGTGAGATCCAGTTTCGATTAACCCACTCGTGCACCACTGATCTTCAGCATC 6960  
Db 6901 ACCGCTGTGAGATCCAGTTTCGATTAACCCACTCGTGCACCACTGATCTTCAGCATC 6960  
Qy 6961 TTTTACTTTTCCACGAGCTTCTGGGTGAGCAAAACAGGAGGCAAAATGCGGCAGAAAAA 7020  
Db 6961 TTTTACTTTTCCACGAGCTTCTGGGTGAGCAAAACAGGAGGCAAAATGCGGCAGAAAAA 7020  
Qy 7021 GGGAAATAGGGCGACACAGGAATGTTGAATACTCATACTCTTCTCTTTTCAATATTATTG 7080  
Db 7021 GGGAAATAGGGCGACACAGGAATGTTGAATACTCATACTCTTCTCTTTTCAATATTATTG 7080  
Qy 7081 AAGCAATTTATCAGGGTTATTTCTCATGAGCGGATACATATTTGAATGATTTTAAAAAAA 7140  
Db 7081 AAGCAATTTATCAGGGTTATTTCTCATGAGCGGATACATATTTGAATGATTTTAAAAAAA 7140  
Qy 7141 TAAACAAATAGGGTTTCCGCGACATTTTCCCGGAAAAAGTGCACCTGACGCTTAAGAAAC 7200  
Db 7141 TAAACAAATAGGGTTTCCGCGACATTTTCCCGGAAAAAGTGCACCTGACGCTTAAGAAAC 7200

Qy 7201 CATTATTATCATGACATTAACTATAAAATAGCGGTATCAAGAGGCCCTTTTCGTCTGCG 7260  
Db 7201 CATTATTATCATGACATTAACTATAAAATAGCGGTATCAAGAGGCCCTTTTCGTCTGCG 7260  
Qy 7261 GCGTTTCGGTGCATGACCGTGAARACCTCTGACACATCGCAGCTCCCGGAGCGGTACACAGC 7320  
Db 7261 GCGTTTCGGTGCATGACCGTGAARACCTCTGACACATCGCAGCTCCCGGAGCGGTACACAGC 7320  
Qy 7321 TTGTCTCTAAGCGGATGCCGGAGCAGACAAGCCCGTCAGGCGCGTCAGCGGGTGTGG 7380  
Db 7321 TTGTCTCTAAGCGGATGCCGGAGCAGACAAGCCCGTCAGGCGCGTCAGCGGGTGTGG 7380  
Qy 7381 CGGTTGTGCGGGCTGGCTTAACTATATGCGGCATCAGACAGATTTGTACTGAGAGTGCACCA 7440  
Db 7381 CGGTTGTGCGGGCTGGCTTAACTATATGCGGCATCAGACAGATTTGTACTGAGAGTGCACCA 7440  
Qy 7441 TATGCGGTGTGAATATACCGCACAGATCGTAAAGGAGAAATACCGCATCAGCGGCCATTTC 7500  
Db 7441 TATGCGGTGTGAATATACCGCACAGATCGTAAAGGAGAAATACCGCATCAGCGGCCATTTC 7500  
Qy 7501 GCCATTGAGGCTGCGCAACTGTTGGGAAGGCGCATCGTGGGCGCTCTTCGCTATTACG 7560  
Db 7501 GCCATTGAGGCTGCGCAACTGTTGGGAAGGCGCATCGTGGGCGCTCTTCGCTATTACG 7560  
Qy 7561 CCAGCTGCGGAAAGGGGGATGTGCTGCAAGCGGATTAAGTTGGGTAAACGCCAGGTTTTTC 7620  
Db 7561 CCAGCTGCGGAAAGGGGGATGTGCTGCAAGCGGATTAAGTTGGGTAAACGCCAGGTTTTTC 7620  
Qy 7621 CCAGTACAGGTTGTTAAACGACGCGCCAGTGCC 7654  
Db 7621 CCAGTACAGGTTGTTAAACGACGCGCCAGTGCC 7654

RESULT 2  
ABX13168  
ID ABX13168 standard; DNA; 7654 BP.  
XX  
AC ABX13168;  
XX  
DT 13-MAY-2003 (first entry)  
XX  
DE Retroviral vector encoding a humanised anti-CEA antibody, hMN14.  
XX  
KW CEA; carcinoembryonic antigen; ds; IgTcR; T-cell receptor; cancer;  
KW tumour; colorectal cancer; breast cancer; lung cancer; hMN14; cytostatic;  
KW mouse; human; zeta signalling chain; CD8alpha hinge; humanised antibody;  
KW retroviral vector.  
XX  
OS Retrovirus sp.  
OS Homo sapiens.  
OS Mus sp.  
OS Synthetic.  
OS Chimeric.  
XX  
Key Location/Qualifiers  
CDS 2428..3759  
FT /\*tag= a  
FT /product= "Chimaeric IgTcR protein"  
XX  
XX US2002165360-A1.  
XX  
XX 07-NOV-2002.  
XX  
XX 10-DEC-2001; 2001US-00006771.  
XX  
XX 30-NOV-2000; 2000US-0250087P.  
XX 30-NOV-2000; 2000US-0250090P.  
XX  
XX (JUNG/) JUNGHANS R P.  
XX  
XX Junghans RP;  
XX





Db 1681 ACCGGATCGCTCACAAACAGTCCGTAGATGTCAAGAGAGAGCGTTGGGTTACCTTCTGCT 1740  
Qy 1741 CTGCGAATATGGCAACCTTTAACTGTCGATGGCCGCGAGAGCGCACCTTTAAACGAGACC 1800  
Db 1741 CTGCGAATATGGCAACCTTTAACTGTCGATGGCCGCGAGAGCGCACCTTTAAACGAGACC 1800  
Qy 1801 TCATCACCCAGGTTAAGATCAAGGTCCTTTTACCTGCGCCGCGATGGACACCCAGACCG 1860  
Db 1801 TCATCACCCAGGTTAAGATCAAGGTCCTTTTACCTGCGCCGCGATGGACACCCAGACCG 1860  
Qy 1861 TCCCTCATCATCGTACCTGGGAAGCCTTGGGCTTTTGAACCCCTCCCTGGGTCAAGCCCT 1920  
Db 1861 TCCCTCATCATCGTACCTGGGAAGCCTTGGGCTTTTGAACCCCTCCCTGGGTCAAGCCCT 1920  
Qy 1921 TTGTACACCCTAAGCCTCCGCTCTCTTCTCTCCATCCGCGCGCTCTCTCCCTTTGAAC 1980  
Db 1921 TTGTACACCCTAAGCCTCCGCTCTCTTCTCTCCATCCGCGCGCTCTCTCCCTTTGAAC 1980  
Qy 1981 CTCCTCGTTGCAACCCGCGCTCGATCCTCGCTTTATCCAGCCCTCACTCCTTCTTAGGCG 2040  
Db 1981 CTCCTCGTTGCAACCCGCGCTCGATCCTCGCTTTATCCAGCCCTCACTCCTTCTTAGGCG 2040  
Qy 2041 CCCCATATGSCATATGAGATCTTATATGSGGCACCCCGCCCTTGTAACTTCCCTG 2100  
Db 2041 CCCCATATGSCATATGAGATCTTATATGSGGCACCCCGCCCTTGTAACTTCCCTG 2100  
Qy 2101 ACCCTGACATGACAAAGGTTACTAAACAGCCCTCTCTCAAGCTCACTTACAGGCTTCTA 2160  
Db 2101 ACCCTGACATGACAAAGGTTACTAAACAGCCCTCTCTCAAGCTCACTTACAGGCTTCTA 2160  
Qy 2161 CTTAGTCAGCAACGAGTCTGAGACCTTCTGCGGCGAGCCTTCAAGAAACAACTGGACCG 2220  
Db 2161 CTTAGTCAGCAACGAGTCTGAGACCTTCTGCGGCGAGCCTTCAAGAAACAACTGGACCG 2220  
Qy 2221 ACCGTTGTAACCTCGCTGAAAGGACCTTACACAGTCTCTGCTGACACCCCGCACAGAC 2280  
Db 2221 ACCGTTGTAACCTCGCTGAAAGGACCTTACACAGTCTCTGCTGACACCCCGCACAGAC 2280  
Qy 2281 TAAGAACCTAGAACCTCGCTGAAAGGACCTTACACAGTCTCTGCTGACACCCCGCACAGC 2340  
Db 2281 TAAGAACCTAGAACCTCGCTGAAAGGACCTTACACAGTCTCTGCTGACACCCCGCACAGC 2340  
Qy 2341 CCTCAAAGTAGACGCGATCGAGCTTGATACACGCGCCCAACGTTGAAGGTCGCGACCC 2400  
Db 2341 CCTCAAAGTAGACGCGATCGAGCTTGATACACGCGCCCAACGTTGAAGGTCGCGACCC 2400  
Qy 2401 CGGGGTTGGAACCTCTAGACTGCCATGGATGGAGTGTATCATCTCTTCTTGTA 2460  
Db 2401 CGGGGTTGGAACCTCTAGACTGCCATGGATGGAGTGTATCATCTCTTCTTGTA 2460  
Qy 2461 GCAACAGCTACAGGTTGCCACTCCGACATCCAGCTGACCCAGAGCCCAAGCAGCTGAGC 2520  
Db 2461 GCAACAGCTACAGGTTGCCACTCCGACATCCAGCTGACCCAGAGCCCAAGCAGCTGAGC 2520  
Qy 2521 GCCAGCTGGGTGACAGAGTACCATCACTGTAAAGGCCAGTCAAGATGTGGTACTTCT 2580  
Db 2521 GCCAGCTGGGTGACAGAGTACCATCACTGTAAAGGCCAGTCAAGATGTGGTACTTCT 2580  
Qy 2581 GTAGTTGGTACCGAGAGCGAGGTAAGGTTCAAGGTTCAAAAGCTGTGATCTA CTGAGCATCC 2640  
Db 2581 GTAGTTGGTACCGAGAGCGAGGTAAGGTTCAAGGTTCAAAAGCTGTGATCTA CTGAGCATCC 2640  
Qy 2641 ACCGGCACACTGGTGTGCCAGAGATTCAGCGGTAGCGGTAGCGGTACCGATTACCC 2700  
Db 2641 ACCGGCACACTGGTGTGCCAGAGATTCAGCGGTAGCGGTAGCGGTAGCGGTACCGATTACCC 2700  
Qy 2701 TTCCACATCAGCAGCCTCCAGCCAGAGACATCGCCACTACTACTGCGGCAATATAGC 2760  
Db 2701 TTCCACATCAGCAGCCTCCAGCCAGAGACATCGCCACTACTACTGCGGCAATATAGC 2760  
Qy 2761 CTCTATCGGTCTGGCCAAAGGACCAAGGTGGAATCAAAACGAGGTGGCTCAGGATCG 2820  
Db 2761 CTCTATCGGTCTGGCCAAAGGACCAAGGTGGAATCAAAACGAGGTGGCTCAGGATCG 2820

Qy 2821 GGTGATCCGCTCTGTGGTCTCAGGATCGAGGTCCAACCTGGTGGAGAGCGGTGAGGT 2880  
Db 2821 GGTGATCCGCTCTGTGGTCTCAGGATCGAGGTCCAACCTGGTGGAGAGCGGTGAGGT 2880  
Qy 2881 GTTGTGAACCTGCGCGCTCTGCGCTGCTCTCGCATCTGCGCTTCTGATTTTCAAC 2940  
Db 2881 GTTGTGAACCTGCGCGCTCTGCGCTGCTCTCGCATCTGCGCTTCTGATTTTCAAC 2940  
Qy 2941 ACATATTGGATGAGTTGGGTGAGACAGGACCTGGAAAAGGTCTTTGAGTTGGATTGGAGAA 3000  
Db 2941 ACATATTGGATGAGTTGGGTGAGACAGGACCTGGAAAAGGTCTTTGAGTTGGATTGGAGAA 3000  
Qy 3001 ATTCAATCAGATAGCAGTACGATTAACTATGCGCGCTCTTAAAGGATAGATTTCATA 3060  
Db 3001 ATTCAATCAGATAGCAGTACGATTAACTATGCGCGCTCTTAAAGGATAGATTTCATA 3060  
Qy 3061 TCGGAGACCAACCGCCAGAACACATTTGTTCTGCAAAATGGACAGCTGAGACCCGAGAC 3120  
Db 3061 TCGGAGACCAACCGCCAGAACACATTTGTTCTGCAAAATGGACAGCTGAGACCCGAGAC 3120  
Qy 3121 ACCGGGTCTATTTTGTGCAAGCCTTCTGCGCTTCCCTGCTGCTTATTTGGGCG 3180  
Db 3121 ACCGGGTCTATTTTGTGCAAGCCTTCTGCGCTTCCCTGCTGCTTATTTGGGCG 3180  
Qy 3181 CAAGGGAACCCCGGTCAACCGTCTCCAGTGTCTAAGGCCACCAAGAGCGCGCGACCA 3240  
Db 3181 CAAGGGAACCCCGGTCAACCGTCTCCAGTGTCTAAGGCCACCAAGAGCGCGCGACCA 3240  
Qy 3241 CCAACACCGCGCCCAACCATCGCTGCGAGCCCTGCTGCGCCAGAGCGCGCTCGG 3300  
Db 3241 CCAACACCGCGCCCAACCATCGCTGCGAGCCCTGCTGCGCCAGAGCGCGCTCGG 3300  
Qy 3301 CCAGCGCGGGGCGGAGTGCACACGAGGGGCTGGAATTCGCCCTGATCCCAAACTC 3360  
Db 3301 CCAGCGCGGGGCGGAGTGCACACGAGGGGCTGGAATTCGCCCTGATCCCAAACTC 3360  
Qy 3361 TGCTACCTGCTGGATGGAATCTCTTCACTATGTTGTCATTTCTCACTGCTTGTTCCTG 3420  
Db 3361 TGCTACCTGCTGGATGGAATCTCTTCACTATGTTGTCATTTCTCACTGCTTGTTCCTG 3420  
Qy 3421 AGATGAGTTTCAGAGAGCGAGAGCCCGCGCTTACAGAGGGGCGAGAACGAGCTC 3480  
Db 3421 AGATGAGTTTCAGAGAGCGAGAGCCCGCGCTTACAGAGGGGCGAGAACGAGCTC 3480  
Qy 3481 TATAACGAGCTCAATCTAGGACGAAGAGAGGATACGATGTTTGGACAAGAGAGCTGCG 3540  
Db 3481 TATAACGAGCTCAATCTAGGACGAAGAGAGGATACGATGTTTGGACAAGAGAGCTGCG 3540  
Qy 3541 CGGACCTCGAGATGGGGGAAAGCCGAGAGGAAAGAACCTCAGGAAAGGCTGTACAAT 3600  
Db 3541 CGGACCTCGAGATGGGGGAAAGCCGAGAGGAAAGAACCTCAGGAAAGGCTGTACAAT 3600  
Qy 3601 GAACTGCAAGATAAGATGCGGAGGCTTACAGTGAAGTTGGATGAAAGGCGAGCGC 3660  
Db 3601 GAACTGCAAGATAAGATGCGGAGGCTTACAGTGAAGTTGGATGAAAGGCGAGCGC 3660  
Qy 3661 CGGAGGCGAAGGGCGAGATGGCTTTACAGGCTCTCAGTACAGCCACCAAGAGACCC 3720  
Db 3661 CGGAGGCGAAGGGCGAGATGGCTTTACAGGCTCTCAGTACAGCCACCAAGAGACCC 3720  
Qy 3721 TACGACCCCTTACATGTCAGGCGCTGCTGCTTAACTCGAGCGCGCGGATCGG 3780  
Db 3721 TACGACCCCTTACATGTCAGGCGCTGCTGCTTAACTCGAGCGCGCGGATCGG 3780  
Qy 3781 GATTAGTCCAAATTTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTGACTCAAC 3840  
Db 3781 GATTAGTCCAAATTTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTGACTCAAC 3840  
Qy 3841 AATATCAGCTGAGGCTTATAGATACGAGCCTATAGATATAAATAAAGATTATTTA 3900  
Db 3841 AATATCAGCTGAGGCTTATAGATACGAGCCTATAGATATAAATAAAGATTATTTA 3900

QY 3901 GTCTCAGAAAAAGGGGGGAATGAAGACCCACCTGTAGGTTTGGCAAGCTAGCTTAAG 3960  
DB 3901 GTCTCAGAAAAAGGGGGGAATGAAGACCCACCTGTAGGTTTGGCAAGCTAGCTTAAG 3960  
QY 3961 TAAAGCCATTTTGCAGAGGCAATGGAATAATCATAACTGAGAAATAGAGAACTCAGATCAAG 4020  
DB 3961 TAAAGCCATTTTGCAGAGGCAATGGAATAATCATAACTGAGAAATAGAGAACTCAGATCAAG 4020  
QY 4021 GTTAGGAAACAGAGAGACAGAGATATGGGCCAAACAGAGATATCTGTGTAGAGATGCTCC 4080  
DB 4021 GTTAGGAAACAGAGAGACAGAGATATGGGCCAAACAGAGATATCTGTGTAGAGATGCTCC 4080  
QY 4081 TGCCCCGCTCAGGGCCAAAGAACAGTTGGAACAGGAGAAATATGGGCCAAACAGAGATATCTG 4140  
DB 4081 TGCCCCGCTCAGGGCCAAAGAACAGTTGGAACAGGAGAAATATGGGCCAAACAGAGATATCTG 4140  
QY 4141 TGGTAAGCAGTTCTCTGCCCGCTCAGGGCCAAAGAACAGATGGTCCCAAGATGCGGTGCC 4200  
DB 4141 TGGTAAGCAGTTCTCTGCCCGCTCAGGGCCAAAGAACAGATGGTCCCAAGATGCGGTGCC 4200  
QY 4201 GCCCTCAGCAGTTCTTAGAGAACCATCAGATGTTCCAGGGTGCCCCAAGGACCTGAAT 4260  
DB 4201 GCCCTCAGCAGTTCTTAGAGAACCATCAGATGTTCCAGGGTGCCCCAAGGACCTGAAT 4260  
QY 4261 GACCCGTGCTGCTTATTTGAATCAACCAATCAGTTGCTTCTCGCTTCTGCTCGCGGCTT 4320  
DB 4261 GACCCGTGCTGCTTATTTGAATCAACCAATCAGTTGCTTCTCGCTTCTGCTCGCGGCTT 4320  
QY 4321 CTGCTCCCCAGTCAATAAAGAGCCCAACACCCCTACTCGGGCGCCAGTCTCCGA 4380  
DB 4321 CTGCTCCCCAGTCAATAAAGAGCCCAACACCCCTACTCGGGCGCCAGTCTCCGA 4380  
QY 4381 TAGACTGCGTCGCGGGGTACCGGTGTTCTCAATAAACCTCTTGAGTGTGATCCGACT 4440  
DB 4381 TAGACTGCGTCGCGGGGTACCGGTGTTCTCAATAAACCTCTTGAGTGTGATCCGACT 4440  
QY 4441 CGTGGTCTCGCTGTTCTTGGGAGGTCTCTCTGAGTGAATGCAAGAGTGTCCCTTCAAGCCAGCT 4500  
DB 4441 CGTGGTCTCGCTGTTCTTGGGAGGTCTCTCTGAGTGAATGCAAGAGTGTCCCTTCAAGCCAGCT 4500  
QY 4501 CTTTTCAGTTTCTCCCACTACACAGGTCTCACTAACATCTCTGATGTGCCGAGGACTC 4560  
DB 4501 CTTTTCAGTTTCTCCCACTACACAGGTCTCACTAACATCTCTGATGTGCCGAGGACTC 4560  
QY 4561 CGTCAGCCCGTTTTTGTATTAATAAATGCAAGAGTGTCCCTTCAAGCCAGCT 4620  
DB 4561 CGTCAGCCCGTTTTTGTATTAATAAATGCAAGAGTGTCCCTTCAAGCCAGCT 4620  
QY 4621 ACATCTGACTCTCGGCTTTATAAAGAAATGTTGAAGGCTCTGTGACATCTTGCCACA 4680  
DB 4621 ACATCTGACTCTCGGCTTTATAAAGAAATGTTGAAGGCTCTGTGACATCTTGCCACA 4680  
QY 4681 CGACTTTTTAAGATTTTTATGCTCTCGATGAGGATTTAGTCAATCTATCCTCGTCTA 4740  
DB 4681 CGACTTTTTAAGATTTTTATGCTCTCGATGAGGATTTAGTCAATCTATCCTCGTCTA 4740  
QY 4741 TTTTGTGCTTCTCGTATTTTAAATTTCTAGTTTGAATCTGCTTCTGAGAGCAGGCT 4800  
DB 4741 TTTTGTGCTTCTCGTATTTTAAATTTCTAGTTTGAATCTGCTTCTGAGAGCAGGCT 4800  
QY 4801 GATTGACAGATGATTAAATCTGAGGGCAGGCTCTGTGAAAGGTTCCCTGGGCTCAG 4860  
DB 4801 GATTGACAGATGATTAAATCTGAGGGCAGGCTCTGTGAAAGGTTCCCTGGGCTCAG 4860  
QY 4861 TGTGAGATTTGCAATAAAGGGGCTCTGCCCCCTGTGTACAGACAGATCGGAATCTAGA 4920  
DB 4861 TGTGAGATTTGCAATAAAGGGGCTCTGCCCCCTGTGTACAGACAGATCGGAATCTAGA 4920  
QY 4921 GTGCATCTCAGAGTCCCGGGTCCGGGCTCTGATCTCAGGGATCTTTGCCCTAGAG 4980  
DB 4921 GTGCATCTCAGAGTCCCGGGTCCGGGCTCTGATCTCAGGGATCTTTGCCCTAGAG 4980  
QY 4981 ATCCTCTACGCGGACGCATCGTGGCGGGTACCGAGCTCGAATTCGTAATCATGTGCAT 5040

DB 4981 ATCCTCTACGCGGACGCATCGTGGCGGGTACCGAGCTCGAATTCGTAATCATGTGCAT 5040  
QY 5041 AGCTGTTTCTGTGTGAATTTTATCCGCTCACAATTTCCACACAATACAGAGCGGAA 5100  
DB 5041 AGCTGTTTCTGTGTGAATTTTATCCGCTCACAATTTCCACACAATACAGAGCGGAA 5100  
QY 5101 GCATAAAGTGTAAAGCCTTGGGTGCTTAATGAGTGAGCTAATCACTCAATTAATTTGGTTGC 5160  
DB 5101 GCATAAAGTGTAAAGCCTTGGGTGCTTAATGAGTGAGCTAATCACTCAATTAATTTGGTTGC 5160  
QY 5161 GCTCACTGCCCCCTTTCCAGTCCGGAAACCTGTCTGTCGAGCTGATTAATGAATCGGCC 5220  
DB 5161 GCTCACTGCCCCCTTTCCAGTCCGGAAACCTGTCTGTCGAGCTGATTAATGAATCGGCC 5220  
QY 5221 AAGCGCGGGGAGAGCGGTTTGGCTATTTGGGCGCTCTTCCGCTTCTCGCTCACTGACT 5280  
DB 5221 AAGCGCGGGGAGAGCGGTTTGGCTATTTGGGCGCTCTTCCGCTTCTCGCTCACTGACT 5280  
QY 5281 CGCTCGCTCGGTCTGCGCTGCGGCGAGGGTATCAGCTCACTCAAGGCGGTAATAC 5340  
DB 5281 CGCTCGCTCGGTCTGCGCTGCGGCGAGGGTATCAGCTCACTCAAGGCGGTAATAC 5340  
QY 5341 GGTATTCACAGAAATCAGGGGATACGACAGGAAAGAACATGTGAGCAAAAGGCCAGCAA 5400  
DB 5341 GGTATTCACAGAAATCAGGGGATACGACAGGAAAGAACATGTGAGCAAAAGGCCAGCAA 5400  
QY 5401 AGGCAGAGAAACCGTAAAAAGGCGGTTGCTGGCGGTTTTTCCATAGGCTCCGCCCCCTG 5460  
DB 5401 AGGCAGAGAAACCGTAAAAAGGCGGTTGCTGGCGGTTTTTCCATAGGCTCCGCCCCCTG 5460  
QY 5461 ACGAGCATCACAAAATCGAGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAA 5520  
DB 5461 ACGAGCATCACAAAATCGAGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAA 5520  
QY 5521 GATACAGCGGTTTCCCTCGGAACTCCCTCGTGGCTCTCTCTGTTCCGACCTCGCGC 5580  
DB 5521 GATACAGCGGTTTCCCTCGGAACTCCCTCGTGGCTCTCTCTGTTCCGACCTCGCGC 5580  
QY 5581 TTACCGGATACCTGTCCGCTTTCTCCCTTGGGAAAGGCTGGCGCTTCTCATAGTCAAC 5640  
DB 5581 TTACCGGATACCTGTCCGCTTTCTCCCTTGGGAAAGGCTGGCGCTTCTCATAGTCAAC 5640  
QY 5641 GCTGTAGGTATCTCAGTTTCGTTAGGTGCTTCCGCTCAAGCTGGCTGTGTGTCAGCAAC 5700  
DB 5641 GCTGTAGGTATCTCAGTTTCGTTAGGTGCTTCCGCTCAAGCTGGCTGTGTGTCAGCAAC 5700  
QY 5701 CCCCCGTTCAGCCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAAACCCGG 5760  
DB 5701 CCCCCGTTCAGCCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAAACCCGG 5760  
QY 5761 TAAAGACAGATTTATTCGCACTGGCAGAGCCACTGGTAAACAGGATTTAGCAGAGGAGT 5820  
DB 5761 TAAAGACAGATTTATTCGCACTGGCAGAGCCACTGGTAAACAGGATTTAGCAGAGGAGT 5820  
QY 5821 ATGTAGGCGGTCTACAGAGTTCTTGAAGTGGTGGCTTAACTACGCTACACTAGAAGA 5880  
DB 5821 ATGTAGGCGGTCTACAGAGTTCTTGAAGTGGTGGCTTAACTACGCTACACTAGAAGA 5880  
QY 5881 CAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCT 5940  
DB 5881 CAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCT 5940  
QY 5941 CTTGATTCGGGCAAAACAAACCCGCTGGTAGCGGTTTGGTTTGGTTCAGAGCAGAGA 6000  
DB 5941 CTTGATTCGGGCAAAACAAACCCGCTGGTAGCGGTTTGGTTTGGTTCAGAGCAGAGA 6000  
QY 6001 TTACCGCGAGAAAAAGGATCTCAAGAGATCTTGTGATCTTTTCTACGGGCTCTGAGC 6060  
DB 6001 TTACCGCGAGAAAAAGGATCTCAAGAGATCTTGTGATCTTTTCTACGGGCTCTGAGC 6060  
QY 6061 CTTAGTGAAACGAAACCTCACTTAAAGGATTTTGGTCAAGAGATTTCAAAAAAGGATCT 6120





QY 5213 AATCGCCAAACGCGCGGGAGAGGCGGTTTGGGTAATTGGGCGCTCTTCCGCTTCCCTCGCT 5272  
DB 7040 AATCGCCAAACGCGCGGGAGAGGCGGTTTGGGTAATTGGGCGCTCTTCCGCTTCCCTCGCT 7099  
QY 5273 CACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 5332  
DB 7100 CACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 7159  
QY 5333 GGTAAATACGGTTATCCACAGAAATCAGGGGATTAACGACAGAAAGAACATGTGAGCAAAAGG 5392  
DB 7160 GGTAAATACGGTTATCCACAGAAATCAGGGGATTAACGACAGAAAGAACATGTGAGCAAAAGG 7219  
QY 5393 CCAGCAAAAGGCGAGGAACCGTTAAAGGCGCGCTGCTGCGGCTTTTCCATAGGCTCCG 5452  
DB 7220 CCAGCAAAAGGCGAGGAACCGTTAAAGGCGCGCTGCTGCGGCTTTTCCATAGGCTCCG 7279  
QY 5453 CCCCCCTCACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAACCAGACAGG 5512  
DB 7280 CCCCCCTCACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAACCAGACAGG 7339  
QY 5513 ACTATAAGATACACGAGGCTTCCCTCGAAGCTCCCTCGTGGCTCTCCTGTTCCGAC 5572  
DB 7340 ACTATAAGATACACGAGGCTTCCCTCGAAGCTCCCTCGTGGCTCTCCTGTTCCGAC 7399  
QY 5573 CCTGCGCTTACCGGATACCTGTCGCTTCTCCCTTCGGGAAAGCGTGGCGCTTTCTCA 5632  
DB 7400 CCTGCGCTTACCGGATACCTGTCGCTTCTCCCTTCGGGAAAGCGTGGCGCTTTCTCA 7459  
QY 5633 TAGCTACGCTGTAGGTATCTCAGTTCGGTGTAGTTCGCTCCAGCTGGGCTGTGT 5692  
DB 7460 ATGCTACGCTGTAGGTATCTCAGTTCGGTGTAGTTCGCTCCAGCTGGGCTGTGT 7519  
QY 5693 GCAGAAACCCCGCTTCAGCCCGACGCTCGGCTTATCCGCTAACTATCTGCTTGGATC 5752  
DB 7520 GCAAGAACCCCGCTTCAGCCCGACGCTCGGCTTATCCGCTAACTATCTGCTTGGATC 7579  
QY 5753 CAACCCGTTAGACACGCTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTAGCAG 5812  
DB 7580 CAACCCGTTAGACACGCTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTAGCAG 7639  
QY 5813 AGCGAGGTATGTAGGCGGTGTACAGAGTCTTGAAGTGGTGGCTAACTACGGGTACAC 5872  
DB 7640 AGCGAGGTATGTAGGCGGTGTACAGAGTCTTGAAGTGGTGGCTAACTACGGGTACAC 7699  
QY 5873 TAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGTTACTTTCGGAAGAGT 5932  
DB 7700 TAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGTTACTTTCGGAAGAGT 7759  
QY 5933 TGGTAGCTCTTGTATCCGGCAAAACAAACCCAGCTGGTAGCGGTGGTTTTTTTGTGCAA 5992  
DB 7760 TGGTAGCTCTTGTATCCGGCAAAACAAACCCAGCTGGTAGCGGTGGTTTTTTTGTGCAA 7819  
QY 5993 GCAGCAGATTAACGCGCAGAAAAAGGATCTCAAGAGATCCTTTGTATCTTCTACGG 6052  
DB 7820 GCAGCAGATTAACGCGCAGAAAAAGGATCTCAAGAGATCCTTTGTATCTTCTACGG 7879  
QY 6053 GTCTGACGCTCAGTGGAAAGAAAACTCAGTTAAGGATTTTGGTCAAGATTAATCAA 6112  
DB 7880 GTCTGACGCTCAGTGGAAAGAAAACTCAGTTAAGGATTTTGGTCAAGATTAATCAA 7939  
QY 6113 AAGGATCTTACCTAGATCTTTTAAATTAATAAATGAAGTTTAAATCAATCTAAAGTAT 6172  
DB 7940 AAGGATCTTACCTAGATCTTTTAAATTAATAAATGAAGTTTAAATCAATCTAAAGTAT 7999  
QY 6173 ATATAGTAAATCTGGTCTGACGTTACCAATGCTTAATAGTAGGCACTTATCTCAGC 6232  
DB 8000 ATATAGTAAATCTGGTCTGACGTTACCAATGCTTAAATAGTAGGCACTTATCTCAGC 8059  
QY 6233 GATCTGCTATTTCTGTTCAATAGTTGCTGACTCCCGCTGCTGTAGATAAATCTAGAT 6292  
DB 8060 GATCTGCTATTTCTGTTCAATAGTTGCTGACTCCCGCTGCTGTAGATAAATCTAGAT 8119

QY 6293 ACGGAGGGCTTACCATCTGSCCCAGTGTCTGCAATGATACCGCGAGACCCACGCTCACC 6352  
DB 8120 ACGGAGGGCTTACCATCTGSCCCAGTGTCTGCAATGATACCGCGAGACCCACGCTCACC 8179  
QY 6353 GGCTCCAGATTTATCAGCAATAAACACAGCAGCGGAGGCGCGAGCGCAAGTGTGTCC 6412  
DB 8180 GGCTCCAGATTTATCAGCAATAAACACAGCAGCGGAGGCGCGAGCGCAAGTGTGTCC 8239  
QY 6413 TGCAACTTTATCCGCCCTCCATCCAGTCTATTAAATTGTTGCCGGGAAGCTAGAGTAAGTAG 6472  
DB 8240 TGCAACTTTATCCGCCCTCCATCCAGTCTATTAAATTGTTGCCGGGAAGCTAGAGTAAGTAG 8299  
QY 6473 TTCCGCCAGTTAATAGTTTGGCAACGTTGTTGCCATTGCTACAGGC-TGCTGGTGTCTACG 6531  
DB 8300 TTCCGCCAGTTAATAGTTTGGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCTACG 8359  
QY 6532 CTGCTGCTTTGCTATGCTTCAATTCAGTCCGGTTCCCAACGATCAAGCGAGTTTACATG 6591  
DB 8360 CTGCTGCTTTGCTATGCTTCAATTCAGTCCGGTTCCCAACGATCAAGCGAGTTTACATG 8419  
QY 6592 ATCCCCCATGTTGTGCAAAAAAGCGTTAGTCTCTTCGCTCCTCGATTCGTTGTCTCAGAG 6651  
DB 8420 ATCCCCCATGTTGTGCAAAAAAGCGTTAGTCTCTTCGCTCCTCGATTCGTTGTCTCAGAG 8479  
QY 6652 TAAGTTGGCCGACGTTATTCATCTAGTGGTATGGCAGCACTGCATATTTCTTTACTGT 6711  
DB 8480 TAAGTTGGCCGACGTTATTCATCTAGTGGTATGGCAGCACTGCATATTTCTTTACTGT 8539  
QY 6712 CATGCCATCCGTAGATGCTTTCTGTGACTGGTGTACTCAACCAAGTCAATTCCTGAGA 6771  
DB 8540 CATGCCATCCGTAGATGCTTTCTGTGACTGGTGTACTCAACCAAGTCAATTCCTGAGA 8599  
QY 6772 ATAGTGTATCGGCGACCGAGTTGCTCTTTCGCCCGCGTCAATAGCGGATAATACCGCGCC 6831  
DB 8600 ATAGTGTATCGGCGACCGAGTTGCTCTTTCGCCCGCGTCAATAGCGGATAATACCGCGCC 8659  
QY 6832 ACATAGCAGAACTTTAAAGTGTCTATCATTTGGAAAAAGTTTCTTCGGGGGAAAACTCTC 6891  
DB 8660 ACATAGCAGAACTTTAAAGTGTCTATCATTTGGAAAAAGTTTCTTCGGGGGAAAACTCTC 8719  
QY 6892 AAGGATCTTACCGCTGTGAGATCCAGTTGATTAACCCACTGTCACCCACTGATC 6951  
DB 8720 AAGGATCTTACCGCTGTGAGATCCAGTTGATTAACCCACTGTCACCCACTGATC 8779  
QY 6952 TTCAAGTCTTTTACTTTTACCCAGCGTTTCTGCGGTGAGCAAAAAACAGGAAGCAAAATGC 7011  
DB 8780 TTCAAGTCTTTTACTTTTACCCAGCGTTTCTGCGGTGAGCAAAAAACAGGAAGCAAAATGC 8839  
QY 7012 CGCAAAAAAGGGAATAAGGCGCACACGGAATGTTGAATACTCATACTCTTTCTTTTCA 7071  
DB 8840 CGCAAAAAAGGGAATAAGGCGCACACGGAATGTTGAATACTCATACTCTTTCTTTTCA 8899  
QY 7072 ATATTATGAAGCATTTTATCAGGTTTATGTTCTCATGAGCGGATACATATTGATGTAT 7131  
DB 8900 ATATTATGAAGCATTTTATCAGGTTTATGTTCTCATGAGCGGATACATATTGATGTAT 8959  
QY 7132 TTAGAAAAATAAACAATAGGCGTTCGCGCACATTTCCCGAAAAAGTGCACCTGAGCT 7191  
DB 8960 TTAGAAAAATAAACAATAGGCGTTCGCGCACATTTCCCGAAAAAGTGCACCTGAGCT 9019  
QY 7192 CTAAAGAAACCAATTATTCATGACATTAACCTATAAAAAATAGGGGTATCACGAGGCCCTT 7251  
DB 9020 CTAAAGAAACCAATTATTCATGACATTAACCTATAAAAAATAGGGGTATCACGAGGCCCTT 9079  
QY 7252 TCCTCTCGCGGTTTCGGTGTATGACGGTGAACCTCTGACACATGACGTCCCGGAGAC 7311  
DB 9080 TCCTCTCGCGGTTTCGGTGTATGACGGTGAACCTCTGACACATGACGTCCCGGAGAC 9139  
QY 7312 GGTCAACAGTGTCTGTAAAGCGGATGCCGGGAGCAGACAAGCCGCTCAGGGCGGTCTCAGC 7371  
DB 9140 GGTCAACAGTGTCTGTAAAGCGGATGCCGGGAGCAGACAAGCCGCTCAGGGCGGTCTCAGC 9199  
QY 7372 GGGTGTGGCGGGTGTGGGGGCTGGCTTAACTATGCGCGCATCAGAGCAGAGATTGTACTGAG 7431





Db 1479 CCCCTACATCGTGACCTGGGAAGCCCTGGCTTTTGACCCCTCTCCCTGGGTCAAGCCCTT 1538  
Qy 1922 TGTCACCCCTAAGCCTCGCCTCTCTCTTCCTCATCCGCCCGCTCTCTCCCTTTGAACC 1981  
Db 1539 TGTCACCCCTAAGCCTCGCCTCTCTCTTCCTCATCCGCCCGCTCTCTCCCTTTGAACC 1598  
Qy 1982 TCCTCGTTGACCCCGCCTCGATCCTCCCTTTATCCAGCCCTCACTCTCTCTAGGGCC 2041  
Db 1599 TCCTCGTTGACCCCGCCTCGATCCTCCCTTTATCCAGCCCTCACTCTCTCTAGGGCC 1658  
Qy 2042 CCCCATATGGCATATGAGATCTTATATGGGGCACCCCGCCCTCTTGTAACCTTCCCTGA 2101  
Db 1659 -----CCATATGAGATCTTATATGGGGCACCCCGCCCTCTTGTAACCTTCCCTGA 1708  
Qy 2102 CCCTGACATGACAAGAGTTACTAACAAGCCCTCTCTCCAAGCTCACTTACAGGCT-TCTA 2160  
Db 1709 CCCTGACATGACAAGAGTTACTAACAAGCCCTCTCTCCAAGCTCACTTACAGGCTCTCTA 1768  
Qy 2161 CTTAGTCCAGACGGAAGTCTGGAGACCTCTGGCGGACCTACCAAGAACCACTGGACCG 2220  
Db 1769 CTTAGTCCAGACGGAAGTCTGGAGACCTCTGGCGGACCTACCAAGAACCACTGGACCG 1828  
Qy 2221 ACCGGTGTACTCTACCCCTTACCGAGTCGGGACACAGTGTGGTCCGCCGACACCGAC 2280  
Db 1829 ACCGGTGTACTCTACCCCTTACCGAGTCGGGACACAGTGTGGTCCGCCGACACCGAC 1888  
Qy 2281 TAAGAACCTAGAACCTCGCTGGAAGGACCTTACACAGTCTCTGTGACACCCGCCACCGC 2340  
Db 1889 TAAGAACCTAGAACCTCGCTGGAAGGACCTTACACAGTCTCTGTGACACCCGCCACCGC 1948  
Qy 2341 CCTCAAAGTAGACGGCATCGAGCTTGATGATACAGCCGCCCACTGTGAAGGCTGGGACCC 2400  
Db 1949 CCTCAAAGTAGACGGCATCGAGCTTGATGATACAGCCGCCCACTGTGAAGGCTGGGACCC 2008  
Qy 2401 CGGGGTGGACCATCTCTAGACTGCCATGCCATGGATGGAGCTGTATCATCTCTTCTGTGA 2460  
Db 2009 CGGGGTGGACCATCTCTAGACTGCCATGCCATGGATGGAGCTGTATCATCTCTTCTGTGA 2068  
Qy 2461 GCAACAGCTACAGGTGTCTCACTCCGACATCCAGCTGACCCAGAGCCCAAGCAGCTGAGC 2520  
Db 2069 GGTACCCGGGGGCGCGCGATCTTATTAAGCGGCGCGGTAAACA----- 2116  
Qy 2521 GCCAGCTGGGTGACAGGTGACCATCACTGTGAAGGCCAGTCAAGATGGGTACTTCT 2580  
Db 2117 -----TTG 2119  
Qy 2581 GTAGCTTGGTACCAGACAGCCAGGTAAGGCTCCAAAGCTGCTGATCTAGTGACATCC 2640  
Db 2120 TTAACCTAACTAAGTAGACAGGTTTCCCTCTAGCGGGATCAATTCCTCCCTCTCCC 2179  
Qy 2641 ACCCGGCACACTGGTGTGCCAAGCAGATTACAGCGGTAGCGGTACCGACTTCAACC 2700  
Db 2180 TCCCTCCCTTAAGCTTACTGGCGA----- 2206  
Qy 2701 TTCACCTACAGAGCTCCAGCCAGAGACATCGCCACTACTACTGCCAGCAATATAGC 2760  
Db 2207 -----AGCCGCTTGGAAATAAGCCGGTGTGGTTGTCTATATGTTATTTTCCACC 2257  
Qy 2761 CTCTATCGTCTGGCCAGGACCAAGGTGGAAATCAACAGGCTGGCTCAGGATCG 2820  
Db 2258 ATATGGCGCTTTTGGCAATGTGA----- 2282  
Qy 2821 GGTGGATCCGCTCTGGTGGCTCAGGATCGGAGGTCCAACTGGTGGAGAGCGGTGAGGT 2880  
Db 2283 -----GGCCCGGAAACCTG 2297  
Qy 2881 GTTGTGCAACTGGCCGGTCCCTGCGCTGCTCTGCTCCGATCTGGCTTCGATTTCAACC 2940  
Db 2298 GCCCTGTCTTCTGACGAGCATCTCTAGGGCTCTTCCCTCTCGCCAAAGGAATGCAAG 2357  
Qy 2941 ACATATTGGATGAGTTGGGTGAGACGACCACTGGAAAGGTCTTGTAGTGGATTTGGAGAA 3000  
Db 2358 GTCTGTTGAATGTCGTGAAGGAAGCATGTTCTCTCTGGAAAGCTTCTTGAAGACAAACAGT 2417

Qy 3001 ATTCAATCCAGATACAGTACGATTAACTATCGCGCGTCTCTTAAAGGATAGATTTACAATA 3060  
Db 2418 CTGTAGCGA-----CCCTTTGACGGCAGCGGAACCCCCACCTGGCGACAGGTGCTCTCGC 2474  
Qy 3061 TCGCGAGACCAACGCCAAGAACACATTTGTTCTGCAATGGACAGCTGAGACCCGGAAC 3120  
Db 2475 GCCAAAAGCCACGCTGATATAGATACACTGCTGAAAGGGGCGCAACCCCGAGTGCACGTTG 2534  
Qy 3121 ACCGGGGTCTATTTTGTGCAAGCCTTTACTTTCGGCTTCCCTGCTGTTTGTCTTATTTGGGCG 3180  
Db 2535 TGAGTTGATAGTTGTGGAAAGAGTCAATGGCTCTCTCAAGCGTATTTCAACAGGGGCG 2594  
Qy 3181 CAAGGGACCCCGGTCAACGCTCTCAGTCTGAAGCCCAACACGACGCCAGCGCCGCGACA 3240  
Db 2595 TGAAGGATGCCACAGAAGGTACCCCAT----- 2621  
Qy 3241 CCAACA CCGGCGCCCA CCAATCGCGTCGCGCCCTGTCTCTGCGCCAGAGGGCGCTCGG 3300  
Db 2622 -----GTATGGG 2628  
Qy 3301 CCAGCGCGGGGCGCGAGTGCACACGAGGGGGCTGGACTTCGCGCTGGATCCCAAACTC 3360  
Db 2629 ATCTGATCTGGGGCTCGGTGCACATGCTTTACATGTGTTTAGTCGAGGTTAAAAACGT 2688  
Qy 3361 TGCTACCTGTGGATGGAATCCTCTTCACTATGCTGTCACTTCTCACTGCTTGTCTCTG 3420  
Db 2689 CTAGGCCCCCGAA-----CCACGGGACGTGGTTTTCTCTTTGMAAAAACACGATAATA 2741  
Qy 3421 AGAGTGAAGTTACGAGGCGCGAGAGCCCGCTGCTACGAGGGGCGCAACACGCTC 3480  
Db 2742 CCATGGTCAAGAAACATTTAAACATTTCTCAACAGATCTAGAATTAGTAAGTAGCGA 2801  
Qy 3481 TATAACGAGCTCAATCTAGGACGAAGAGAGAGTACGATGTTTTGGCAAGAGAGCTGCGC 3540  
Db 2802 CAGAGAAGTTACATGCTTTATGAGGATAATAAACATCATGTGGAGCGCAATTCGTA 2861  
Qy 3541 CGGACCCCTGAGATGGGGGGAAGCCGAGAGGAAGAACCTCTCAGGAGGCTCTGTACAAT 3600  
Db 2862 CGAAAACAGGAGAAATCATTTTCGCGCATACATATTTGAAGCGTATATAGGACGAGTAACTG 2921  
Qy 3601 GAATCGCAGAAAGATAGATGGCGGAGCGCTTACAGTGAAGTTGGGATGAAGGGGAGCGC 3660  
Db 2922 TTTGTGAGAAGCAATTCGATTTGGTAG--TGCAGTTTCGAATGGCAAAAAGGATTTTGA 2979  
Qy 3661 CGAGGGGCAAGGGCGCAGATGGCTTTTACAGGGTCTCAGTACAGCCACCAAGACACC 3720  
Db 2980 CACGATTTGAGCTGTTAGACACCCCTTATTTGACGAAGTAGATAGAGTATTCGAGTGGT 3039  
Qy 3721 TACGACCCCTTCACTGCGAGCCCTGCGCTAACTCGACCGCGCGCGGATCCG 3780  
Db 3040 AAGTCTTTGGTATGTGTAG-----GAGTTGATTTTCAAGTATGACCA 3085  
Qy 3781 GATTAGTCCAAATTTGTTAAAGACAGGATATCAGTGGTCTCAGGCTCTAGTTTGTACTAAC 3840  
Db 3086 GATTGTTTGTGTTAAATAGAAATGAATGCGAAGTAGTCAAAACTACGATTTGAAGAATC 3145  
Qy 3841 AATATCACCAAGCTGAAGCTTATAGAGTACGAGCCATAGATAAAATAAAGATTTTATTTA 3900  
Db 3146 ATTCCACTCAATATATCC--GAATTTAAAGTCGACGATAAAATAAAGATTTTATTTA 3203  
Qy 3901 GTCTCCAGAAAAAGGGGGGAATGAAAGACCCCACTCTGTAGGTTTGGCAAGCTAGCTTAAG 3960  
Db 3204 GTCTCCAGAAAAAGGGGGGAATGAAAGACCCCACTCTGTAGGTTTGGCAAGCTAGCTTAAG 3263  
Qy 3961 TAAGCCCATTTTGAAGGCTAGG--AAAAATACATACTGAGATAGAGAGTTAGATCAA 4019  
Db 3264 TAAGCCCATTTTGAAGGCTAGGAAAAATACATACTGAGATAGAGAGTTAGATCAA 3323  
Qy 4020 GGTTAGGAAACAGA-GACACAGCAGAAATAGGGCCAAACAGGATATCTGTGTAGCAGT 4078  
Db 3324 GGTCAAGAAACAGATGGAAACAGCTGAAATAGGGCCAAACAGGATATCTGTGTAGCAGT 3383



Db 5544 AATCAATCTAAAGTATATATAGTAAACTTTGGTCTGACAGTTTACCAATGCTTAATCAGTG 5603  
Qy 6217 AGGCACCTATCTCAGCGATCTGTCTATTTCGTTTCCATCCATAGTCCGTCGACCTCCCGTCG 6276  
Db 5604 AGGCACCTATCTCAGCGATCTGTCTATTTCGTTTCCATCCATAGTCCGTCGACCTCCCGTCG 5663  
Qy 6277 TGTAGATTAACATACGATACGGAGGCTTACCATCTGGCCCAAGTCTGCAATGATACCGC 6336  
Db 5664 TGTAGATTAACATACGATACGGAGGCTTACCATCTGGCCCAAGTCTGCAATGATACCGC 5723  
Qy 6337 GAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGCGCG 6396  
Db 5724 GAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGCGCG 5783  
Qy 6397 AGCGCAGAGTGTCTCGCACTTTATCGGCTCCATCCAGTCTATTAATTTGTTCCGGG 6456  
Db 5784 AGCGCAGCGGTCTCGCACTTTATCGGCTCCATCCAGTCTATTAATTTGTTCCGGG 5843  
Qy 6457 AAGCTAGAGTAGTGTGCGCAGTTAATAGTTTGGCAAGTGTGTCATTGCTACAG 6516  
Db 5844 AAGCTAGAGTAGTGTGCGCAGTTAATAGTTTGGCAAGTGTGTCATTGCTACAG 5903  
Qy 6517 GC-TGCTGGTGTACGCTCGTCTGTTTGGTATGCTTTCATTGAGTCCGCTTCCCAACGAT 6575  
Db 5904 GCATCGTGGTGTACGCTCGTCTGTTTGGTATGCTTTCATTGAGTCCGCTTCCCAACGAT 5963  
Qy 6576 CAAGCGAGTTACATGATCCCGCATGTTGTGCAAAAAGCGTTAGCTCCTTCGCTCCTC 6635  
Db 5964 CAAGCGAGTTACATGATCCCGCATGTTGTGCAAAAAGCGTTAGCTCCTTCGCTCCTC 6023  
Qy 6636 CGATCGTGTGCAAGTAGTGTGCGCGCAGTGTATCATCTCATGCTTATGCGAGCAGCTGC 6695  
Db 6024 CGATCGTGTGCAAGTAGTGTGCGCGCAGTGTATCATCTCATGCTTATGCGAGCAGCTGC 6083  
Qy 6696 ATAATTCCTTACTCTCATGCTCATCCGTAAGATGCTTTTCTGTGACTGCTGACTCA 6755  
Db 6084 ATAATTCCTTACTCTCATGCTCATCCGTAAGATGCTTTTCTGTGACTGCTGACTCA 6143  
Qy 6756 CCAAGTCACTCTGAGAATAGTGTATGCGCGCAGCGAGTGTCTTTCGCGCGGCTCAATAC 6815  
Db 6144 CCAAGTCACTCTGAGAATAGTGTATGCGCGCAGCGAGTGTCTTTCGCGCGGCTCAATAC 6203  
Qy 6816 GGGATAATACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCTTGGAAAAAGCTTCTT 6875  
Db 6204 GGGATAATACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCTTGGAAAAAGCTTCTT 6263  
Qy 6876 CGGCGGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTC 6935  
Db 6264 CGGCGGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTC 6323  
Qy 6936 GTGCACCCAACTGATCTTTCAGCATCTTTTACTTTTCCACAGCGTTTCTGGGTGAGCAAAA 6995  
Db 6324 GTGCACCCAACTGATCTTTCAGCATCTTTTACTTTTCCACAGCGTTTCTGGGTGAGCAAAA 6383  
Qy 6996 CAGGAAGCAAAATGCGCAAAAAGGAATAAGGCGCAGCAGGAATGTTGAATACTCA 7055  
Db 6384 CAGGAAGCAAAATGCGCAAAAAGGAATAAGGCGCAGCAGGAATGTTGAATACTCA 6443  
Qy 7056 TACTCTTCTTTTCAATATTATTTGAACATTTATCAGGGTATGCTCATGAGCGAT 7115  
Db 6444 TACTCTTCTTTTCAATATTATTTGAACATTTATCAGGGTATGCTCATGAGCGAT 6503  
Qy 7116 ACATATTGAATGTATTTAGAAAAATAAACAATAGGGTTTCCGCGCATCTTCCCGAA 7175  
Db 6504 ACATATTGAATGTATTTAGAAAAATAAACAATAGGGTTTCCGCGCATCTTCCCGAA 6563  
Qy 7176 AAGTGCCACCTGACGCTCTAAGAAACCAATTATTATCATGACATTAACCTTAAAAATAGGC 7235  
Db 6564 AAGTGCCACCTGACGCTCTAAGAAACCAATTATTATCATGACATTAACCTTAAAAATAGGC 6623  
Qy 7236 GTATCAGAGGCGCTTTCGTC 7256  
Db 6624 GTATCAGAGGCGCTTTCGTC 6644

## RESULT 5

ADA12886  
ID ADA12886 standard; DNA; 6505 BP.  
XX  
AC ADA12886;  
XX  
DT 20-NOV-2003 (first entry)  
XX Murine MSCV retrovirus vector sequence.  
DE  
DE  
KW ds; virus vector; insulator; gypsy; cHS4 dimer core sequence;  
KW viral regulatory control element; HSC1; stem cell; antianaemic;  
KW anisickling; cytosatic; antiparkinsonian; nootropic; neuroprotective;  
KW gene therapy; thalassaemia; sickle cell anaemia; leukaemia; malignancy;  
KW neurodegenerative disease; Parkinson's disease; Alzheimer's disease;  
KW multiple sclerosis; mouse.  
XX  
OS Mus sp.  
XX  
FN WO2003070958-A1.  
XX  
PD 28-AUG-2003.  
XX  
PF 19-FEB-2003; 2003WO-CA000229.  
XX  
PR 19-FEB-2002; 2002CA-02370841.  
PR 23-FEB-2002; 2002US-0358933P.  
XX  
PA (HOSP-) HOSPITAL FOR SICK CHILDREN.  
XX  
PI Yao S, Ellis J;  
XX  
XX WPI; 2003-663855/62.  
XX  
XX New nucleic acid molecule for a virus vector with silencer-blocking  
insulator activity, for stem cell marking, stem cell genetic manipulation  
studies, gene therapy, or treating thalassemia, sickle cell anemia or  
leukemia.  
XX  
PS Example 1; Page 51-55; 64pp; English.  
XX  
XX The invention relates to a novel nucleic acid for a virus vector with  
silencer-blocking insulator activity that integrates into the genome of a  
mammalian stem cell comprising an insulator element, a viral regulatory  
control element, and a coding nucleic acid molecule operatively  
associated with the viral regulatory element and capable of expression in  
the cell. The insulator element comprises a gypsy sequence, a cHS4 dimer  
core sequence, or all or part of a fully defined sequence of 611 or 722  
base pairs given in the specification. The viral regulatory control  
element comprises an infectious retrovirus vector sequence or a  
lentivirus vector sequence. The infectious retrovirus sequence comprises  
HSC1 with a fully defined sequence of 6143 bp, given in the  
specification. The coding nucleic acid molecule is a reporter gene, where  
the reporter gene is PGK eGFP. The host cell is a stem cell of embryonic  
or adult tissue origin. The nucleic acid of the invention has  
antianaemic, anisickling, cytosatic, antiparkinsonian, nootropic, and  
neuroprotective activity. The vector may have a use in gene therapy. The  
nucleic acid is useful for stem cell marking or stem cell genetic  
manipulation studies. The nucleic acid or vectors and compositions  
comprising it are also useful in stem cell gene therapy, or for treating  
patients with diseases, disorders or abnormal physical states of stem  
cells, e.g. thalassaemia, sickle cell anaemia, leukaemia, malignancies,  
or neurodegenerative diseases such as Parkinson's disease, Alzheimer's  
disease or multiple sclerosis. Host cells comprising the nucleic acid are  
useful as research tools to measure levels of expression of the coding  
nucleic acid molecule and the activity of the polypeptide encoded by the  
coding nucleic acid molecule. The present sequence represents the murine  
MSCV retrovirus vector sequence.  
XX  
SQ Sequence 6505 BP; 1382 A; 1732 C; 1666 G; 1464 T; 0 U; 261 Other;

	Query Match	42.4%; Score 3244.4; DB 9; Length 6505;	
	Best Local Similarity	71.1%; Pred. No. 0;	
	Matches 5064; Conservative	0; Mismatches 1020; Indels 1035; Gaps	25;
Qy	552	TGAAGACCCACCTGTAGGTTTGGCAAGCTAGCTTAAGTAAGCCAATTTTGCAAGGCAT	611
Db	1	TGAAGACCCACCTGTAGGTTTGGCAAGCTAGCTTAAGTAAGCCAATTTTGCAAGGCAT	60
Qy	612	GGAATAACATAACTAGAAATAGAGAAGTTTCAGATCAAGGTTTAGGAACAAGAGACAGCA	671
Db	61	GGAATAACATAACTAGAAATAGAGAAGTTTCAGATCAAGGTTTAGGAACAAGAGACAGC-	119
Qy	672	GAATATGGGCCAACAGGATATCTGTGGTAAGCAGTTCTGCCCGCTCAGGGCCAGAA	731
Db	120	-----	119
Qy	732	CAGTTGGAACAGGAGNATATGGGCCAACAGGATATCTGTGGTAAGCAGTTCTGCCCG	791
Db	120	-----AGAATATGGGCCAACAGGATATCTGTGGTAAGCAGTTCTGCCCG	166
Qy	792	GCTCAGGGCCAAAGAACAGATGGTCCCAGATGGGTCCCCCTCCAGAGTTCTTAGAGA	851
Db	167	GCTCAGGGCCAAAGAACAGATGGTCCCAGATGGGTCCCCCTCCAGAGTTCTTAGAGA	226
Qy	852	ACCATCAGATGTTTTCCAGGTTGCCCAAAGACCTTG-AAATGACCTGTGCTTATTGAA	910
Db	227	ACCATCAGATGTTTTCCAGGTTGCCCAAAGACCTGAAATGACCTGTGCTTATTGAA	286
Qy	911	CTAACCAATCAGTTTCGCTTCCTGTTCTGTTTCGGGGCTTCGTCCCCGAGTCAATAA	970
Db	287	CTAACCAATCAGTTTCGCTTCCTGTTCTGTTTCGGGGCTTCGTCCCCGAGTCAATAA	346
Qy	971	AAGAGCCCAACAAACCCCTCACCTCGSGCGCCAGTCTCCGATAGACTGCGTGCOCGGETA	1030
Db	347	AAGAGCCCAACAAACCCCTCACCTCGSGCGCCAGTCTCCGATAGACTGCGTGCOCGGETA	406
Qy	1031	CCCGTATTCCEAATAAAGCCCTTTGCTGTTTTGCAATCCGAATCGTGGACTCGCTCATCCTT	1090
Db	407	CCCGTATTCCEAATAAAGCCCTTTGCTGTTTTGCAATCCGAATCGTGGACTCGCTCATCCTT	466
Qy	1091	GGAGGGTCTCTCAGATTGATTGACTTGCCCACTCGGGGCTCTTTCATTGAGGTTCC	1150
Db	467	GGAGGGTCTCTCAGATTGATTGACTTGCCCACTCGGGGCTCTTTCATTGAGGTTCC	526
Qy	1151	ACCGAGATTTGGAGACCCCTGCCAGGGAACAACGACCCCGCCGGAGGTAAAGCTGG	1210
Db	527	ACCGAGATTTGGAGACCCCTGCCTAGGGAACAACGACCCCGCCGGAGGTAAAGCTGG	586
Qy	1211	CCAGCAACTTATCTGTGCTGTGTC-----CGATGCTGTAGTGTCATGACTGATTTTA	1262
Db	587	CCAGCGTCTGTTGCTGCTGTCTCTCTCTTTGTGCGTGTGTGCGGGCATCTAATGTT	646
Qy	1263	TGGCCTTCGCTCGGTACTAGTTAGCTAACTAGCTCTGTATCTGCGGACCGGTGGGAA	1322
Db	647	TGGCCTTCGCTCTGTACTAGTTAGCTAACTAGCTCTGTATCTGCGGACCGGTGGGAA	706
Qy	1323	CTGACAGTTTCGGAAACCCCGCCGCAACCTCGGAGACGTCCACAGGACTTCGGGGGCC	1382
Db	707	CTGACAGTTTCGAAACCCCGCCGCAACCTCGGAGACGTCCACAGGACTTCGGGGGCC	766
Qy	1383	GTTTTTGTGGCCCGACTGATGCTCTAAAATCCCGATCGTTTAGGACTCTTTTGGTGCACCC	1442
Db	767	GTTTTTGTGGCCCGACTGAGGAGGGAGTCGATGTGAAATCCGACCC-----	814
Qy	1443	CCCTTAGAGGAGGNATATGTTGTTCTGGTAGGAGACGAGAACCTTAAACAGTTCCCGCCT	1502
Db	815	-----CGTCAGGATATGTTGTTCTGGTAGGAGACGAGAACCTTAAACAGTTCCCGCCT	867
Qy	1503	CCGTCTCAATTTTTGCTTCGGTTTGGGACCGAAGCCGCGCGCTCTGT---CTGC	1559
Db	868	CCGTCTCAATTTTTGCTTCGGTTTGGAACCGAAGCCGCGCTCTGTCTGCTGACGGC	927
Qy	1560	TGCAGCATCGTCTGTGTTGTCCTGTGCTGACTGTGTTCTGTATTTGTCTGAAAAATATG	1619



Db 1929 CGACCTGACGCAATATGATGGATCGCCATTTGAACAGATGGATTGACGCGAGTTCTC-- 1986  
Qy 2700 CTTCCACATCAGACGCTCCAGCCAGAGGACATCGCCACTACTACTCTGACGCAATATAG 2759  
Db 1987 -----CGGCCGCTTTGGGTGGAGAGGCTAT--TCGGCTATGACTGGGCACACAGAC 2035  
Qy 2760 CCTCTATCGTCTGCTCGCCAAAGGACCAAGGTGGAATCAACAGAGTGCTCAGGATC 2819  
Db 2036 AATCGCTGCTC-----TGATCGCGCGGT 2062  
Qy 2820 GGGTGGATCCGCGCTCTGCTGCTCAGGATCGAGTCCAACTGCTGGAGAGCGGTGGAGG 2879  
Db 2063 CCGGCTGTACGACGAGGGGCGCCCGGTTCTTTTGTCAAGACCGACTGTCCGGTGCCT 2122  
Qy 2880 TGTGTGCAACTGCGCGCTCCCTGCGCTCTGCTGCTCGCATCTGCTTCGATTTTCAC 2939  
Db 2123 GAATGAACCTGCAGGACGAGGACGCGGCTATCGTGGCTGCGCCACGAGCGGCTTCTTG 2182  
Qy 2940 CACATATTTGGATGAGTTGGGTGAGACAGGCACCTGGAAAAAGGTCTTGAGTGGAATGGAGA 2999  
Db 2183 GCGAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGACTGGCTGTCTATTTGGGCGAAGT 2242  
Qy 3000 AATTCATCCAGATGACGATGATTAATCTATGCGCC-GTCTCTAAAGGATAGATTACAA 3058  
Db 2243 GCGGGGCGAGATCTCTGTCTATCTCACTTGTCTCTGCCGAGAAAGTATCCATCATGGC 2302  
Qy 3059 TATCGCAGACAAACGCCAAGACACATTTGTTCTGCAATGACAGCGCTGAGACCCGAAG 3118  
Db 2303 TGATGCANTCGGCGCTGCATACGCTTGATCCGGCTACCTGCCATTCGACCCACGAC 2362  
Qy 3119 ACACGGGGTCTATTTTGTGGAAGCTTTACTTGGCTTCCCTGCTTCCCTGCTTATTTGGG 3178  
Db 2363 GAAAC---ATGCATCGAGCGACGCTACTCGGATGGAAGCGCTTGTGATCAGGA 2419  
Qy 3179 GCCAAGGACCCCGCTCAGCTCTCCAGTGTAGCCCAACAGACGCGCCAGCGCGGAC 3238  
Db 2420 TGATCTGGAAGAGCATCAGGGGCTCGCGCCACGCGCAATGTTTCGCCAGGCTCAAGGC 2479  
Qy 3239 CACCAACACCGCGCGCCACCATCGGCTCGAGCCCTGCTGCTGCGCCAGAGGGGGCTC 3298  
Db 2480 GCGCATGCCAGCGCGAGGATCTGCTGTAACCATGGC----- 2519  
Qy 3299 GCGCAGCGCGGGGGCGCAGTGCAACAGAGGGGCTGAGCTTCGCGCTCGATCCCAAC 3358  
Db 2520 -----GATGCTGCT 2529  
Qy 3359 TCTGCTACTGCTGATGGAATCCTCTTCATCTATGTTGATCTACTGCTTGTTC 3418  
Db 2530 TGCCGAATATCATGTTGGAATGCGCGCTTTTCTGGAATTCATCGACTGTG----- 2580  
Qy 3419 TGAGAGTGAAGTTACAGGAGCGCAGAGCCCCCGGTACAGCAGGCGCCAGAACACGAC 3478  
Db 2581 -----GCGGCTGGGTGTGGGAC 2601  
Qy 3479 TCTATAACGAGCTCAATCTAGGACGAAGAGGAGTACGATGTTTGGACGAAGACGCTG 3538  
Db 2602 GCTATCAG----- 2610  
Qy 3539 GCGGGGACCTGAGATGGGGGGAAGCGCAGAGGAAGAACCTCAGGAAGGCTGTACA 3598  
Db 2611 ----- 2610  
Qy 3599 ATGAATGCAGAAAGATAGATGGCGAGGCTACAGTGAGATGGGATGAAGCGGAC 3658  
Db 2611 -----ACATAGGCTGGCTACCCGCTGATGATGCTGAAGAGCTTGGCG 2652  
Qy 3659 GCGGAGGGGCAAGGGGCGATGCGCTTTTACAGGGTCTCAGTACAGCCACCAAGGACA 3718  
Db 2653 GGAATGGGTGACCGCTTCTCTGCTTTAGGGTATCGCCGCTCCCGATGGCAGCGCA 2712  
Qy 3719 CCTACGACGCCCTTCAATGACAGGCCCTGCCCCCTCGCTTAACGACGCGCGCGGATC 3778  
Db 2713 TCGCCTTCTATCGCCTTCTTGAC----- 2736

Qy 3779 CGGATTAAGTCCAATTTGTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTGTACTCA 3838  
Db 2737 -----AGTTCTTCTGAGGGATCGTCCGACCTGCGAGCAAGCTT----- 2775  
Qy 3839 ACAATATCACCGCTGAAGCCTATAGAGTACGAGCCATAGATAAAATAAAAGATTTTATT 3898  
Db 2776 -----ATCGATAAAATAAAAGATTTTATT 2799  
Qy 3899 TAGTCTCCAGAAAAGGGGGAATGAAAGACCCACCTGTAGGTTTGGCAAGCTAGCTTA 3958  
Db 2800 TAGTCTCCAGAAAAGGGGGAATGAAAGACCCACCTGTAGGTTTGGCAAGCTAGCTTA 2859  
Qy 3959 AGTAAACCCCATTTTGAAGCATGGAATAATACATACTGAGAAATGAGAGATTGAGATCA 4018  
Db 2860 AGTAAACCCCATTTTGAAGCATGGAATAATACATACTGAGAAATGAGAGATTGAGATCA 2919  
Qy 4019 AGGTTAGGAACAGAGACAGCAGAAATATGGCCAAACAGGATATCTGTGTGAAGCAGTT 4078  
Db 2920 AGGTTAGGAACAGAGACAGC----- 2941  
Qy 4079 CTTGCCCCGCTCAGGGGCAAGAACAGTTGGAAACAGGAGAAATATGGGCCAAACAGGATATC 4138  
Db 2942 -----AGAAATATGGGCCAAACAGGATATC 2965  
Qy 4139 TGTGTAAGCAGTTCTTGCCTCGGCTCAGGCGCAAGAACAGATGGTCCCAGATGCGGTC 4198  
Db 2966 TGTGTAAGCAGTTCTTGCCTCGGCTCAGGCGCAAGAACAGATGGTCCCAGATGCGGTC 3025  
Qy 4199 CCGCCCTCAGCAGTTCTTAGAGAACCATCAGATGTTTCCAGGGTGCCCAAGGACCTGAA 4257  
Db 3026 CCGCCCTCAGCAGTTCTTAGAGAACCATCAGATGTTTCCAGGGTGCCCAAGGACCTGAA 3085  
Qy 4258 AATGACCTGTGCTTATTTGAACCAATCAGTTCTGCTTCTGCTTCTGTTGCGCG 4317  
Db 3086 AATGACCTGTGCTTATTTGAACCAATCAGTTCTGCTTCTGCTTCTGTTGCGCG 3145  
Qy 4318 CTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGCGGCCAGTCTCTC 4377  
Db 3146 CTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGCGGCCAGTCTCTC 3205  
Qy 4378 CGATAGACTGGTCCGCGGGTACCGGTGTTCTCAATAAACCCTCTTGCAGTTGCATCCG 4437  
Db 3206 CGATAGACTGGTCCGCGGGTACCGGTGTTCTCAATAAACCCTCTTGCAGTTGCATCCG 3265  
Qy 4438 ACTCGTGTCTGCTGTTCTTGGAGGGTCT-CTCTGAGTGTGACTACCCGTCAGCG 4496  
Db 3266 ACTGTGCTGCTGCTGTTCTTGGAGGGTCTCTCTGAGTGTGACTACCCGTCAGCG 3325  
Qy 4497 GGGTCTTTTCAGTTTCTCCACCTACACAGGCTCTCACTAAACATCTCTGATGTGCCGAGGG 4556  
Db 3326 GGGGTCTT----- 3333  
Qy 4557 ACTCCGTCAGCCCGGTTTTTGTATATAAATGCAAGAACAGTGTTCCTTCAAGCCA 4616  
Db 3334 ----- 3333  
Qy 4617 GACTACATCTGACTCTCGGCTTTATAAAGAAATGTTGAAGGGCTCTGTGGACTATCTGC 4676  
Db 3334 ----- 3333  
Qy 4677 CACAGACTTTTAAAGATTTTATGCTCTCGATGAGGATTTAGTCAATCTATCTCTCG 4736  
Db 3334 ----- 3333  
Qy 4737 TCTATTTGTGGCTTCTCGGATTTTAAATTTCTAGTTGCACTCCCTTCTCGAGACA 4796  
Db 3334 ----- 3333  
Qy 4797 CGCGATTGACAGTAGTTAATCTCTGAGGGCAGGCTTCTGTGAAAGGTTGCTGCGGC 4856  
Db 3334 ----- 3333



Db 5462 AAAAGGGAATTAAGGCGACACGGAATGTTGAATCTACTACTCTCTCTTTTCAATAT 5521  
Qy 7076 TATTGAAGCATTTATCAGGGTTATTGTCTCATGACGGATACATATTTGAATGATTTAG 7135  
Db 5522 TATTGAAGCATTTATCAGGGTTATTGTCTCATGACGGATACATATTTGAATGATTTAG 5581  
Qy 7136 AAAATAAACAATAGGGGTTCCGGCGACATTTCCCGAAAAAGTGCACCTGACGTCTAA 7195  
Db 5582 AAAATAAACAATAGGGGTTCCGGCGACATTTCCCGAAAAAGTGCACCTGACGTCTAA 5641  
Qy 7196 GAAACATTTATTATCATGACATTAACCTTATAAATAAGCGGTATCAACGAGCGCTTTTGGT 7255  
Db 5642 GAAACATTTATTATCATGACATTAACCTTATAAATAAGCGGTATCAACGAGCGCTTTTGGT 5701  
Qy 7256 CTCGGCGGTTTCGGTGATGACGGTGAAACCTCTGACATGACACTCCCGGAGACGGTC 7315  
Db 5702 CTCGGCGGTTTCGGTGATGACGGTGAAACCTCTGACATGACACTCCCGGAGACGGTC 5761  
Qy 7316 ACAGTTTGTCTGAAGCGGATGCCGGGAGCAGACAAAGCCGTCAGGGCGGTCAGCGGGT 7375  
Db 5762 ACAGTTTGTCTGAAGCGGATGCCGGGAGCAGACAAAGCCGTCAGGGCGGTCAGCGGGT 5821  
Qy 7376 GTTGGCGGGTTCGGGGTGGCTTAACCTATGCGGCATCAGACGATTTGTACTGAGTG 7435  
Db 5822 GTTGGCGGGTTCGGGGTGGCTTAACCTATGCGGCATCAGACGATTTGTACTGAGTG 5881  
Qy 7436 CACCATATGCGGTGTAATACCGCACAGATGCGTAAGGAGAAATACCGATCAGGGCG 7495  
Db 5882 CACCATATGCGGTGTAATACCGCACAGATGCGTAAGGAGAAATACCGATCAGGGCG 5941  
Qy 7496 CATTGCGCATTCAGGCTCGCAACTGTTTGGGAAGGCGGATCGGTGCGGGCTCTTCGCTA 7555  
Db 5942 CATTGCGCATTCAGGCTCGCAACTGTTTGGGAAGGCGGATCGGTGCGGGCTCTTCGCTA 6001  
Qy 7556 TTAGCCAGCTGCGCAAGGGGGATGCTGCAAGCGCATTAAGTTGGTTACGCCAGG 7615  
Db 6002 TTAGCCAGCTGCGCAAGGGGGATGCTGCAAGCGCATTAAGTTGGTTACGCCAGG 6061  
Qy 7616 TTTTCCAGTCAGACGTTGTAACACGACGCGCAGTGCC 7654  
Db 6062 TTTTCCAGTCAGACGTTGTAACACGACGCGCAGTGCC 6100

RESULT 6  
ID ADQ80674  
XX ADQ80674 standard; DNA; 5903 BP.  
AC ADQ80674;  
XX  
DT 21-OCT-2004 (first entry)  
XX  
DE Bcl-XL expression directing retroviral vector, pFB-bcl-XL.  
XX  
KW Survival; neuron; tyrosine hydroxylase; tyrosine 3-monooxygenase; TH;  
KW anti-apoptotic; Bcl-XL; neurological disorder; neuroprotective;  
KW nootropic; antiparkinsonian; transplantation; drug screening;  
KW gene profiling; CNS disorder; neurodegenerative disease; pFB-bcl-XL; ds.  
XX  
OS Unidentified.  
XX  
XX WO2004062554-A2.  
XX  
XX 29-JUL-2004.  
XX  
XX 07-JAN-2004; 2004WO-DK000008.  
XX  
XX 08-JAN-2003; 2003US-0438719P.  
PR 11-APR-2003; 2003DK-00000581.  
PR 22-APR-2003; 2003US-0464546P.  
XX  
XX (NSGE-) NSGENE AS.  
XX  
XX Martinez-Serrano A, Liste I, Villa A;

XX WPI; 2004-544027/52.  
XX  
XX Enhancing the survival of neurons or cells expressing tyrosine  
PT hydroxylase (TH) for treating neurodegenerative disorders, comprises  
PT contacting neurons or TH expressing cells with Bcl-XL or its functional  
PT equivalent.  
XX  
XX Disclosure; SEQ ID NO 5; 108pp; English.  
XX  
XX The invention relates to a novel method for enhancing the survival of  
CC neurons and/or of cells expressing tyrosine hydroxylase (EC 1.14.16.2 -  
CC tyrosine 3-monooxygenase) (TH + ). The method comprises contacting a  
CC population of cells with Bcl-XL or its functional equivalent, where the  
CC population of cells is selected from: neurons or cells capable of  
CC differentiating into neurons; or TH expressing cells or cells capable of  
CC differentiating into TH expressing cells. The invention further  
CC comprises: a composition of cells obtainable by the method above; a  
CC composition of isolated mammalian cells overexpressing the anti-apoptotic  
CC Bcl-XL protein; a neural progenitor cell; a differentiated dopaminergic  
CC neuron; an implantable cell culture device comprising: a semi-permeable  
CC membrane permitting the diffusion of a biologically active protein  
CC through it; and a composition of cells selected from above; a lentiviral  
CC vector particle being produced based on a lentiviral transfer vector;  
CC enhancing the survival of TH + cells in vivo; a retroviral particle being  
CC produced based on a retroviral transfer vector; enhancing the survival of  
CC in vivo differentiated dopaminergic neurons; a packaging cell line  
CC capable of producing an infective vector particle; a packaging cell line  
CC capable of producing an infective vector particle; treatment of a  
CC neurological disorder; a fusion protein comprising the Bcl-XL sequence  
CC comprising 233 amino acids ADQ80670 or its functional equivalent and a  
CC membrane translocation signal; an expression vector comprising a  
CC polynucleotide sequence coding for the fusion protein and a promoter  
CC sequence capable of directing the expression of the fusion protein in a  
CC host cell; a host cell comprising the expression vector; and producing  
CC the fusion protein. The compositions of the invention have  
CC neuroprotective, nootropic, and antiparkinsonian activities. The cells  
CC are useful for transplantation, drug screening, gene profiling, or for  
CC the preparation of a medicament useful for the treatment of a CNS  
CC disorder. The CNS disorder is a neurodegenerative disease involving  
CC lesioned and traumatic neurons, including traumatic lesions of peripheral  
CC nerves, the medulla, the spinal chord, cerebral ischaemic neuronal  
CC damage, neuropathy, peripheral neuropathy, Alzheimer's disease,  
CC Huntington's disease, Parkinson's disease, amyotrophic lateral sclerosis,  
CC or memory impairment connected to dementia. The method is useful for  
CC enhancing the survival of neurons and/or of cells expressing tyrosine  
CC hydroxylase for the treatment of neurodegenerative disorders. This  
CC sequence represents an expression construct for directing the expression  
CC of Bcl-XL in transduced cells, used in the method for increasing the  
CC survival rate of neurons of the invention.  
XX  
XX Sequence 5903 BP; 1410 A; 1621 C; 1507 G; 1365 T; 0 U; 0 Other;  
SQ

Query Match 42.1%; Score 3223.6; DB 13; Length 5903;  
Best Local Similarity 73.9%; Pred. No. 0;  
Matches 4964; Conservative 0; Mismatches 759; Indels 995; Gaps 29;  
Qy 550 TTTGAAGACCCACCTGTAGTGTGGCAAGCTAGCTTAAGTAAGCCATTTTGAAGGC 609  
Db 167 TGTGAAGACCCACCTGTAGTGTGGCAAGCTAGCTTAAGTAAGCCATTTTGAAGGC 226  
Qy 610 ATGG-AAAATACATACTGAGAATAGAGAAGTTCAGATCAAGGTTAGGAACAGA-GAGAC 667  
Db 227 ATGGAATAATACATACTGAGAATAGAGAAGTTCAGATCAAGGTCAGGAACAGATGGAC 286  
Qy 668 AGCAGAATATGGGCCAAACAGGATATCTGTGTGAAGCAGTTCTCTGCCCC-GCTCAGGGCC 726  
Db 287 AGCTGAATATGGGCCAAACAGGATATCTGTGTGAAGCAGTTCTCTGCCCCGGCTCAGGGCC 346  
Qy 727 AAGAACAGTTTGAACAGGAGAAATATGGGCCAAACAGGATATCTGTGTGAAGCAGTTCTCTG 786  
Db 347 AAGAACAGATGGAACAGCAGCTGAATATGGGCCAAACAGGATATCTGTGTGAAGCAGTTCTCTG 406



Db 2460 ATAAACCCAGGAGCAGATATATCAGAGCTTTGAAACAGGTAGTG----- 2501  
 Qy 3005 ATCCAGATAGCAGTACGATTAACTATATGCGCGTCTCTAAAGGATAGATTACAAATATCGC 3064  
 Db 2502 ----- 2501  
 Qy 3065 GAGACACGCCAAGAAACACATTGTTCCCTGCAAAATGACAGCCTGAGACCCCGAAGACACCG 3124  
 Db 2502 -----AATGAACCTCTTTTCGGGATGGGGTAAACTG 2530  
 Qy 3125 GGGTCAATTTTGTGCAAGCCTTTACTTCCGGCTTCCCTGGTTTCTGTTATTGGGGCCAAG 3184  
 Db 2531 GGGTCGAATGT----- 2542  
 Qy 3185 GGACCCCGGTCAACGCTCTCAGTGTCTAAGCCACACAGCGCGCGGACCAACCA 3244  
 Db 2543 ----- 2542  
 Qy 3245 CACGGGCGCCACATCGGTGCGAGCCCTGTCTCCCTGGCCGACAGAGCGGTGCGCCAG 3304  
 Db 2543 -----GGCCTCTCTCTTTGGGGGGGCACTGTGGTGGAAAGCGTAGACAAGGAGATGC 2598  
 Qy 3305 CGGCGGGGGCGCAGTGCACAGAGGGGCTGGACTTCGCCCTGATCCCAAACTCTGCT 3364  
 Db 2599 AGGTATTGGTGAATGCGAATGCAAGTTGGAATGGCCACCTACTGAAATGACCACTAGAGC 2658  
 Qy 3365 ACCTGCTGGATGGAACTCTCTTCACTATGATGTGTCAATTCCTCACTGCCCTTGTCTCTGAGAG 3424  
 Db 2659 CTTGGATCCAGGAGAACGGC-----GGCTGGGACACTTTTG 2694  
 Qy 3425 TGAAGTTGAGAGGCGCAGAGCCCCCGGTACCAAGAGGGCGAGAACAGCTCTATA 3484  
 Db 2695 TGGATCTCTAGCGGAAACAAATACAGCACCCGAGAGCCGGAAGGCGCAGGAGGTTTCAACC 2754  
 Qy 3485 ACGAGCTCAATCTAGGACGAGAGAGGAGTACGATGTTTTGGACAGAGAGCTGGCCGGG 3544  
 Db 2755 GCTGCTTCTGACGGGCGATGCTGTGGCTGTGTAGTTCTG----- 2795  
 Qy 3545 ACCCTGAGATGGGGGAAAGCCGAGAGGAAGAACCCCTCAGGAAGCCCTGTACAATGAAC 3604  
 Db 2796 ----- 2795  
 Qy 3605 TGCAGAAAGATAAGATGGCGAGGCTTACAGTGAATGGGATGAAGGCGAGCGCCGA 3664  
 Db 2796 ----- 2795  
 Qy 3665 GGGCAGGGGSCAGATGGCCTTTACCAGGGTCTCAGTACAGCCACCAGACACCTTAG 3724  
 Db 2796 ----- 2795  
 Qy 3725 ACGCCCTTCACATGCAAGCCCTGCCCCCTCGCTAACTCGACGCGCGCGGATCGGATT 3784  
 Db 2796 -----CTGGGCTCACTCTTCACTCGGAGTGAACAGATT 2829  
 Qy 3785 AGTCAATTTGTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTTGACTCAACAATA 3844  
 Db 2830 CTGCAGATATCCATCACACTGGCGCGGAATTCGGGGCGCGCATCCGAGTCTTCTGA 2889  
 Qy 3845 TCACCGAGTGAAGCCTATAGAGTAGGACCATAGATAAAATAAAAGATTTTATTAGTCT 3904  
 Db 2890 GCG-----GGACTCTGGGGTTCGATAAAATAAAAGATTTTATTAGTCT 2933  
 Qy 3905 CCAGAAAAGGGGGAATGAAGACCCCACTGTAGGTTTCGGCAAGCTAGCTTAAGTAAC 3964  
 Db 2934 CCAGAAAAGGGGGAATGAAGACCCCACTGTAGGTTTCGGCAAGCTAGCTTAAGTAAC 2993  
 Qy 3965 GCCATTTTGAAGGATGG-AAAAATACATAACTGAGAAATAGAGAAGTTTCAGATCAAGGTT 4023  
 Db 2994 GCCATTTTGAAGGATGGAAAATAATACATAACTGAGAAATAGAGAAGTTTCAGATCAAGGTC 3053  
 Qy 4024 AGGAACAGA-GAGACAGCAGAAATATGGGCCAAAACAGGAATATCTGTGGTAAGCAGTTCTCTG 4082  
 Db 3054 AGGAACAGATGGAACAGCTGAATATGGGCCAAAACAGGATATCTGTGGTAAGCAGTTCTCTG 3113

Qy 4083 CCCC-GCTCAGGGCCAGAACACAGTTGGAAACAGGAGATATATGGGCCAAACAGGATATCTGT 4141  
 Db 3114 CCCCAGCTCAGGGCCAGAACACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGT 3173  
 Qy 4142 GGTAAAGCAGTTCTGCCCCCGCTCAGGGCCAAAGAACAGATGTTGCCAGAGTCGGTCCCG 4201  
 Db 3174 GGTAAAGCAGTTCTGCCCCCGCTCAGGGCCAAAGAACAGATGTTGCCAGAGTCGGTCCCG 3233  
 Qy 4202 CCCTCAGCAGTTCTTAGAGAACCATCAGATGTTTCCAGGGTCCGCCAAAGGACCTGAAATG 4261  
 Db 3234 CCCTCAGCAGTTCTTAGAGAACCATCAGATGTTTCCAGGGTCCGCCAAAGGACCTGAAATG 3293  
 Qy 4262 ACCCTGTGCTTATTTTGAACCAATCAGTTGCTTCTGCTTCTGCTGCGGCTTC 4321  
 Db 3294 ACCCTGTGCTTATTTGAACCAATCAGTTGCTTCTGCTTCTGCTGCGGCTTC 3353  
 Qy 4322 TGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGCGGCGCAGTCTCCGAT 4381  
 Db 3354 TGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGCGGCGCAGTCTCCGAT 3413  
 Qy 4382 AGACTGGTCCCGGGTACCCGTTCTCAATAAACCCCTCTGAGTTGCATCCGACTC 4441  
 Db 3414 TGACTGAGTCCCGGGTACCCGTTCTCAATAAACCCCTCTGAGTTGCATCCGACTC 3473  
 Qy 4442 GTGGTCTGCTGTTCTTGGGAGGCTCT-CTCTGAGTGAATGACTACCCGTCAGCGGGT 4500  
 Db 3474 GTGGTCTGCTGTTCTTGGGAGGCTCTCTCTGAGTGAATGACTACCCGTCAGCGGG 3533  
 Qy 4501 CTTTCAGTTTCTCCCACTACACAGGTTCTCAATAATCTCTGATGTCGCGGAGGACTC 4560  
 Db 3534 TCTTTCAATTTGGGGCTGCTCGGGATCGGGAGACCCCTGCGGAGGACCAAGCCGAC 3593  
 Qy 4561 CGTCAGCCCGTCTTCTGTTATAATAAATCAAGACAGTGTCTTCCCTTCAAGCCAGACT 4620  
 Db 3594 CACCGGAGGTAAGCTG----- 3610  
 Qy 4621 ACATCTGACTCTCGGCTTTATAAAGAAATGTTGAAGGCTCTGTGGACTATCTGCCACA 4680  
 Db 3611 -----GCTGCTCGCGCTTTTCGGTGATGACGGTGAACCTCTG-ACACATGACACTCCC 3665  
 Qy 4681 CGATTTTAAAGATTTTATGCTCTCGGATGAGGGATTTAGTCAATCTATCTCTGCTA 4740  
 Db 3666 GGAGACGCTCACAGTCTGTCTGTAAGCGATGCGGGAGCAGACAAGCCGTCAGGGCGC 3725  
 Qy 4741 TTTTGTGCTCTCTCGTATTTTAAATTTCTAGTTTGCACCTCCCTCTCTGAGACACGCG 4800  
 Db 3726 GTCAGCGGTGTTGGCG----- 3742  
 Qy 4801 GATTGCAGAGTAGTTAATACTCTGAGGCGAGGCTTCTGTGAAGAGTTGCTGCGGCTCAG 4860  
 Db 3743 ----- 3742  
 Qy 4861 TGTGAGATTTTGCATAAAAAGGGTCTGCCCCCTGTGTACAGACAGATCGGAATCTAGA 4920  
 Db 3743 ----- 3742  
 Qy 4921 GTGCATACTCAGAGTCCCCGGGTTCCGGGCTCTGATCTCAGGGCATCTTTGCTAGAG 4980  
 Db 3743 -----GGTGTCCGGGCGCGCCATGA----- 3763  
 Qy 4981 ATCCTACGCCGACGATCGTGGCGGGTACCGAGCTCGAAATTCGTAATCATGGTTCAT 5040  
 Db 3764 ----- 3763  
 Qy 5041 AGCTGTTTCTGTGTGAATTTGTTATCCGCTCAAAATTTCCACAAACATACGAGCCGAA 5100  
 Db 3764 -----CCAGTCAAGTACGATAGCGA 3786  
 Qy 5101 GCATAAAGTAAAGCCTCGGGTCCCTAATGAGTGAAGTGAAGTCACTCACTAATTTGGGTTGC 5160  
 Db 3787 GTGTATATGCTGCTAACTATGCGGCATCAGAGCAGATTGTACTGAGACT----- 3835



QY 5161 GCTACTGCCGCTTTTCAGTCGGGAAACCTGTGTCGCCAGCTGCATTAATGAATCGGCC 5220  
 Db 3836 -----GCACCATATGCGGTGTGAATAACCGC 3861  
 QY 5221 AACCGCGGGAGAGCGGGTTTGGGTATTGGGGGCTCTTCGCTTCTCGCTCACTGACT 5280  
 Db 3862 ACAGATGCGTAAAGGAGAAATACCGCATCAGGCGCTCTTCGCTTCTCGCTCACTGACT 3921  
 QY 5281 CGCTGCGCTCGGTCTGCGTGCAGCGAGCGGTATCAGTCACTCAAAAGGGGTAATAC 5340  
 Db 3922 CGCTGCGCTCGGTCTGCGTGCAGCGAGCGGTATCAGTCACTCAAAAGGGGTAATAC 3981  
 QY 5341 GGTATTCCACAGAAATCAGGGGTAACCGCAGGAAAGAAATGTGAGCAAAAGGCGACAAA 5400  
 Db 3982 GGTATTCCACAGAAATCAGGGGTAACCGCAGGAAAGAAATGTGAGCAAAAGGCGACAAA 4041  
 QY 5401 AGGCCAGGAAACCGTAAAGGCGCGTGTGTCGGCTTTTTCATAGGCTCCGCGCCCTG 5460  
 Db 4042 AGGCCAGGAAACCGTAAAGGCGCGTGTGTCGGCTTTTTCATAGGCTCCGCGCCCTG 4101  
 QY 5461 ACAGCATCACAAAATCAGCGCTCAAGTCAGAGTGGCGAAACCGCAGGACTATAA 5520  
 Db 4102 ACAGCATCACAAAATCAGCGCTCAAGTCAGAGTGGCGAAACCGCAGGACTATAA 4161  
 QY 5521 GATACAGCGGTTTCCCGCTTGAAGCTCCCTGTCGCTCTCTGTTCCGACCCCTGCCGC 5580  
 Db 4162 GATACAGCGGTTTCCCGCTTGAAGCTCCCTGTCGCTCTCTGTTCCGACCCCTGCCGC 4221  
 QY 5581 TTACCGGATACCTGTCGCTTTTCTCCCTTGGGAAAGCGTGGCGTTTCTCATAGTCTAC 5640  
 Db 4222 TTACCGGATACCTGTCGCTTTTCTCCCTTGGGAAAGCGTGGCGTTTCTCATAGTCTAC 4281  
 QY 5641 GCTGTAGGTATCTCAGTTCGGGTAGGTGTCGCTCCAGCTGGGCTGTGTGCGAGAAC 5700  
 Db 4282 GCTGTAGGTATCTCAGTTCGGGTAGGTGTCGCTCCAGCTGGGCTGTGTGCGAGAAC 4341  
 QY 5701 CCCCCCTTCCAGCCGAGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAAACCCGG 5760  
 Db 4342 CCCCCCTTCCAGCCGAGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAAACCCGG 4401  
 QY 5761 TAAGACAGACTTATCGCACATGCGAGCGACGACCTGTTAAAGGATTAAGAGCGAGGT 5820  
 Db 4402 TAAGACAGACTTATCGCACATGCGAGCGACGACCTGTTAAAGGATTAAGAGCGAGGT 4461  
 QY 5821 ATGTAGCGGTGCTACAGAGTCTTCAAGTGTGGCTTAACCTACGCTACCTAGAGGA 5880  
 Db 4462 ATGTAGCGGTGCTACAGAGTCTTCAAGTGTGGCTTAACCTACGCTACCTAGAGGA 4521  
 QY 5881 CAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTCCGAAAAAGAGTTGGTAGCT 5940  
 Db 4522 CAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTCCGAAAAAGAGTTGGTAGCT 4581  
 QY 5941 CTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTTTTGTTCGAGAGCAGA 6000  
 Db 4582 CTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTTTTGTTCGAGAGCAGA 4641  
 QY 6001 TTACGGCGAGAAAAAGGATCTCAAGAGATCTTTTGTATCTTTCTACGGGCTGAGC 6060  
 Db 4642 TTACGGCGAGAAAAAGGATCTCAAGAGATCTTTTGTATCTTTTCTACGGGCTGAGC 4701  
 QY 6061 CTTAGTGGAAACGAAACTCACGTTAAGGATTTTGGTCATGAGATTATCAAAAAGGATCT 6120  
 Db 4702 CTTAGTGGAAACGAAACTCACGTTAAGGATTTTGGTCATGAGATTATCAAAAAGGATCT 4761  
 QY 6121 TCACCTAGATCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATAGT 6180  
 Db 4762 TCACCTAGATCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATAGT 4821  
 QY 6181 AACTTTGGTCTGACGTACCAATGCTTAATCAGTCAGGACCTATCTCAGCGATCTGTC 6240  
 Db 4822 AACTTTGGTCTGACGTACCAATGCTTAATCAGTCAGGACCTATCTCAGCGATCTGTC 4881  
 QY 6241 TATTTCTGTTTCCATGATGTTGCTGACTCCCGCTCGTGTAGATAAATACGAGGAGG 6300

Db 4882 TATTTCTGTTTCCATAGTTGCTGACTCCCGCTGTTAGATAAATACGATACGGAGG 4941  
 QY 6301 GCTTACCATCTGGCCCCAGTGTCAATGATACCGGAGACCCACGCTCACGGCTCCAG 6360  
 Db 4942 GCTTACCATCTGGCCCCAGTGTCAATGATACCGGAGACCCACGCTCACGGCTCCAG 5001  
 QY 6361 ATTATATCAGCAATAAACACAGCAGCGGAAAGGCGCGAGCGAGAGTGGTCTGCAACTT 6420  
 Db 5002 ATTATATCAGCAATAAACACAGCAGCGGAAAGGCGCGAGCGAGAGTGGTCTGCAACTT 5061  
 QY 6421 TATCCGCTTCCATCCAGTCTTAAATTTGTCGGGAAAGCTAGAGTAAGTAGTTCCGCA 6480  
 Db 5062 TATCCGCTTCCATCCAGTCTTAAATTTGTCGGGAAAGCTAGAGTAAGTAGTTCCGCA 5121  
 QY 6481 TTAATAGTTTGGCAACCGTTGTCATTGCTACAGGC-TCCGTGGTGTACCGCTCGTCT 6539  
 Db 5122 TTAATAGTTTGGCAACCGTTGTCATTGCTGCAAGGCATCGTGGTGTACCGCTCGTCT 5181  
 QY 6540 TTGGTATGGCTTTCATTCAGCTCCGGTTCCTCAACGATCAAGCGAGTTACATGATCCCCA 6599  
 Db 5182 TTGGTATGGCTTTCATTCAGCTCCGGTTCCTCAACGATCAAGCGAGTTACATGATCCCCA 5241  
 QY 6600 TGTGTGCAAAAAAGCGGTTAGCTCTTCCGCTCCTCGATCGTTGTCAAGAGTAAAGTTGG 6659  
 Db 5242 TGTGTGCAAAAAAGCGGTTAGCTCTTCCGCTCCTCGATCGTTGTCAAGAGTAAAGTTGG 5301  
 QY 6660 CCGCAGTGTATCAGTCACTGTTATGGCAGCACTGCATAAATCTCTTACTGTGATGCCAT 6719  
 Db 5302 CCGCAGTGTATCAGTCACTGTTATGGCAGCACTGCATAAATCTCTTACTGTGATGCCAT 5361  
 QY 6720 CCGTAAAGTGTCTTTCTGTGACTGTGTGACTCAACCAAGTCATCTCGAGAAATAGTGA 6779  
 Db 5362 CCGTAAAGTGTCTTTCTGTGACTGTGTGACTCAACCAAGTCATCTCGAGAAATAGTGA 5421  
 QY 6780 TCCGCGCACCGAGTTGCTTTTCCCGCGGTCAATACGGGATTAATACCGCGCACATAGCA 6839  
 Db 5422 TCCGCGCACCGAGTTGCTTTTCCCGCGGTCAACACGGGATTAATACCGCGCACATAGCA 5481  
 QY 6840 GAACTTTAAAGTGTCTCATCATGAGAAACGTTCTTCCGGGGGAAACCTCTCAAGATCT 6899  
 Db 5482 GAACTTTAAAGTGTCTCATCATGAGAAACGTTCTTCCGGGGGAAACCTCTCAAGATCT 5541  
 QY 6900 TACCCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGACCCCAACTGATCTTCAGCAT 6959  
 Db 5542 TACCCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGACCCCAACTGATCTTCAGCAT 5601  
 QY 6960 CTTTACTTTTCCACCGCTTCTGGTGAGCAAAAAACAGGAAAGGCAAAATGCCGCAAAA 7019  
 Db 5602 CTTTACTTTTCCACCGCTTCTGGTGAGCAAAAAACAGGAAAGGCAAAATGCCGCAAAA 5661  
 QY 7020 AGGGAATAAGGGGACACGGAATGTTGAATACTCATACTCTTCTCTTTTCAATATTAT 7079  
 Db 5662 AGGGAATAAGGGGACACGGAATGTTGAATACTCATACTCTTCTCTTTTCAATATTAT 5721  
 QY 7080 GAAGCATTTATCAGGTTTATGCTCATGAGCGGATACATATTTGAATGTATTAGAAAA 7139  
 Db 5722 GAAGCATTTATCAGGTTTATGCTCATGAGCGGATACATATTTGAATGTATTAGAAAA 5781  
 QY 7140 ATAAAACAATAGGGGTTCCGCGACATTTTCCCGAAAAAGTGCACCTGACGTCTAAGAAA 7199  
 Db 5782 ATAAAACAATAGGGGTTCCGCGACATTTTCCCGAAAAAGTGCACCTGACGTCTAAGAAA 5841  
 QY 7200 CCAATTATATCATGATTAACCTATATAAATAGGCGTATCAGGAGGCCCTTTCGCT 7257  
 Db 5842 CCAATTATATCATGATTAACCTATATAAATAGGCGTATCAGGAGGCCCTTTCGCT 5899

RESULT 7  
 ABR85887  
 ID . ABR85887 standard; DNA; 5782 BP.  
 XX  
 AC ABR85887;



Db	1317	CATCGTGGAGCCTTGGCTTTTGAACCCCTCCCTGGGTCAAGCCCTTTGTACA	1376	Db	2360	TGTTTTCCGGCACCAAGAGCGGTCCGGAAAGCTGGCTGGAGTGGCAT	2407
Qy	1928	CCCTAAGCCCTCGGCTCTCTTCTTCATCCGCGCCGCTCTCCCGCTTGAACCTCTCG	1987	Qy	3007	CCAGATAGCAGTACGATTAACTATATGCGCCGCTCTCTAAAGGATAGATTACAAATATCGCGA	3066
Db	1377	CCCTAAGCCCTCGGCTCTCTTCTTCATCCGCGCCGCTCTCCCGCTTGAACCTCTCG	1436	Db	2408	--CTTCTTGAGGCCGATATCTGTGCTGCCCTCAACTGGCAGATGACGGTTAGATG	2465
Qy	1988	TTTCGACCCGCTCGATCTCCCTTTATCCAGCCCTCACTCTCTCTAGGGCCGCCCAT	2047	Qy	3067	GACAAAGCGCAGAAACATTTGCTGCAAAATGAGCAGCCTGAGACCCGAAAGACACGGG	3126
Db	1437	TTTCGACCCGCTCGATCTCCCTTTATCCAGCCCTCACTCTCTCTAGGGCCGCCCAT	1496	Db	2466	CGCCCATCTACACCAACGTGACCTATCCATTCGGTCAATCCGCGCTTTGTTCCACGG	2525
Qy	2048	ATGGCATAATGAGATCTTATATGGGACACCCCGCCCTTGTAAACTTCCCTGACCCCTGA	2107	Qy	3127	GTCTATTTTTGFGCAAGCCTTACTTTCGGCTTCCCTGGTTCTTATTTGGGGCCAAAGG	3186
Db	1497	ATGGCATAATGAGATCTTATATGGGACACCCCGCCCTTGTAAACTTCCCTGACCCCTGA	1556	Db	2526	AGAAATCCGACGGGTTGTACTCGCTCACATTTTATTTGATGAAAGCTGGCTACAGGAA	2585
Qy	2108	CATGACAAGAGTTACTAACAAGCCCTCTCTCAAGCTCACTTACAGGCTCTCTACTTAGT	2166	Qy	3187	ACCCCGGTCAACGCTCTCCAGTGAAGCCCAACAGCAGCCAGCCGCGACCAACAACA	3246
Db	1557	CATGACAAGAGTTACTAACAAGCCCTCTCTCAAGCTCACTTACAGGCTCTCTACTTAGT	1616	Db	2586	GGCCAGACCGCAATTTATTTTGTAGTGGCTTAACCTCGGCGTTTCTCTG	2633
Qy	2167	CCAGCACGAAGTCTGGAGACCTCTGGCGCAGCCTACCAAGAACAACTGGACCGACCGGT	2226	Qy	3247	CGCGCGCCCAACCATCGCGTCGAGCCCTGTCTCTGCGGCCAGAGCGGCTCGGCCAGCG	3306
Db	1617	CCAGCACGAAGTCTGGAGACCTCTGGCGCAGCCTACCAAGAACAACTGGACCGACCGGT	1676	Db	2634	-----TGTTGCAACGGGCGCTGGGTGGTACGGGCAAGACAGTCTG	2674
Qy	2227	GGTACCTCAACCTTACCGAGTCGGGACACAGTGTGGGTCCGCGACACAGACTAAGAA	2286	Qy	3307	GGGGGGGCGAGTGCACACAGAGGGGCTGGACTTTCGCCCTGGATCCCAACTCTGTCTAC	3366
Db	1677	GGTACCTCAACCTTACCGAGTCGGGACACAGTGTGGGTCCGCGACACAGACTAAGAA	1736	Db	2675	TTTGGCGCTTAAATTTGAGCTCGAGCGCATATCTACGCGCCGGA---GAAAAACCGCTC	2730
Qy	2287	CCTAGAACCTCGTGGTGAAGACCTTACACAGTCTCTGTGACCAACCCACCGCCCTCAA	1796	Qy	3367	CTGCTGGATGGAATCCTCTTCACTATATGTGTCTCTCACTGCCTTGTCTCTGAGAGTG	3426
Db	1737	CCTAGAACCTCGTGGTGAAGACCTTACACAGTCTCTGTGACCAACCCACCGCCCTCAA	1856	Db	2731	GGGTGTAGTGGTCTCGGCTGGAGTGACGGGAGTTATCT	2768
Qy	2347	AGTAGAGCAGTGGAGCTTGATACACGCCGCCACGTAAGGCTGCCGACCCCGGGGG	2406	Qy	3427	AACTTTCAGCAGGAGCGCAGAGCCCGCGTACACAGAGGGCCAGAACACAGCTCTATAAC	3486
Db	1797	AGTAGAGCAGTGGAGCTTGATACACGCCGCCACGTAAGGCTGCCGACCCCGGGGG	1856	Db	2769	-----TGAATCAAGATATGTGGCGGATGAGCGGATTCGAGCGCAAAAC	2814
Qy	2407	TGGACCATCTCTAGACTGGCATGGAGCTGTATCATCTCTTCTTGTGTAGCAACA	2466	Qy	3487	GAGCTCAATCTAGAGCAGAGAGAGTACGATGTTTTGGACAAGAGAGCTGGCGGGAC	3546
Db	1857	TGGACCATCTCTAGACTGGCATGGAGCTGTATCATCTCTTCTTGTGTAGCAACA	1916	Db	2815	GGTCTGCGCTGGCGACGCGCAATGAAATATATGGCCCAACACAGAGTGGGCGCGCGAC	2874
Qy	2467	GCTACAGGTGTCACTCCGACATCCAGCTGACCCAGAGCCCAAGAGCCTGAGCGCCAGC	2526	Qy	3547	CCTGAGATGGGGGAAGCCGAGAGGAGAACCTCTCAGGAAGGCTGTACATGAACCTG	3606
Db	1917	TATAATCTCGCGCAACTATTTTCCCTCGAAACATTTTAAAGCCGTAGATAAACAGGCT	1976	Db	2875	TTCCAGTT-----CAACATCAGCCGTACAGTCAACAGCAACTGATGGAACACGACCAT	2928
Qy	2527	GTGGGTGACAGAGTACCATCACTGTAGGCCAGTACAGATGTGGGTACTTCTGTAGCT	2586	Qy	3607	CAGAAAGATAGATCGCGAGGCTTACAGTGCAGATTTGGGATGAAAGCGGAGCGCGAGG	3666
Db	1977	GGGACACTTCAATGAGCGAAATAACATCGTCACTGGGACATGTTGACATCATCGC	2036	Db	2929	CGCCATCTCTGCTGCAACGGGAAGAACCGACATGCTGTATACGACGTTTCCATATGGGG	2988
Qy	2587	TGTTACACGACAGGACGAGTGAAGCTCCAAAGCTGCTGATCTACTTGGACATCCACCCGG	2646	Qy	3667	GGCAAGGGGACAGATGGCTTTTACAGGGTCTCAGTACAGCCACCAAGAGACACTACGAC	3726
Db	2037	ACGTAAACTCGCAAGCCGACTGATGC-----CTTCTGACAAATGGAAGGCATTTATTG	2089	Db	2989	ATTGGTGGCGACACTCTCTGGAGCCCGTCAAGTATCGGCGGAAATCCAGCTGAGCGCGGT	3048
Qy	2647	CACACTGTGTGCAAGCAGATTCAGCGGTAGCGGTAGCGGTACGAGTTCACCTTACCC	2706	Qy	3727	GCCCTTCACTCAGCGCCCTGCCCTCGCTAACTCGACGCGCGCGGATCCGGATTAG	3786
Db	2090	CCGTAAAGCGGTGCGGCTCTGGTACCGGTGGGTGAAGACCAAGAACAGCACCTCGATC---	2146	Db	3049	CGCTACCATTTACAGTGTGGTCTGAGTCAAAAATATAATAAACCGGGGAGGCCATCTCTG	3108
Qy	2707	ATCAGAGCCTCCAGCAGGACATCGCACTACTACTCTGCCAGCAATATAGCTCTAT	2766	Qy	3787	TCCAAATTTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTGTACTCAACAATATC	3846
Db	2147	-TGAGCCGCGATATTGCCAGCGGTTTCAACGCGCTGTATGGCGAGATCGATCCCGTC---	2202	Db	3109	CCCGTATTTTCGGGTAGGAAATCCATTATGT-----ACTATTTA	3147
Qy	2767	CGGTCTTTCGGCAGGACCAAGGTGGAAATCAAAACGAGGTGGCTCAGGATCGGGTGA	2826	Qy	3847	ACCAGCTGAAGCCTATAGAGTACGAGCCCATAGATAAAATAAAGATTTTATTAGTCTCC	3906
Db	2203	-----GTTTTACAGCTCGTACCTGGGAAACCCCTGGCGTTA	2239	Db	3148	AACTCGAGGGCGCGCAGCACAGTGGTTCGACGATAAAATAAAGATTTTATTAGTCTCC	3207
Qy	2827	TCCGGCTCTGTGGCTCAGATCGGAGGTCACACTCGTGGAGAGCGGTGAGGTGTGTG	2886	Qy	3907	AGAAAAAGGGGAATGAAGACCCACCTGTAGTTTTCGCAAGCTAGCTTAAGTAACGC	3966
Db	2240	CCAACTTAATGGCTTGGAGACATCCCGCTTTCGCCAGCTGGCGCTTAATAGCAAGAGG	2299	Db	3208	AGAAAAAGGGGAATGAAGACCCACCTGTAGTTTTCGCAAGCTAGCTTAAGTAACGC	3267
Qy	2887	CAACCTGGCGGTCTCGGCTCTCTGCTCCGATCTGGCTTCCATTTCCACCATAT	2946	Qy	3967	CAATTTGCAAGCATGG-AAAAATACATAACTGAGAATAGAGAAGTTTTCAGATCAAGTTAG	4025
Db	2300	CCGCAACCGATCGCCCTTCCCAACAGTTGCGAGCCTGAAATGAGCGAAATGGCGCTTTGCC	2359	Db	3268	CAATTTGGAAGCATGGAATAATACATACTGAGAATAGAGAAGTTTTCAGATCAAGTCTAG	3327
Qy	2947	TGGATGAGTTGGGTGAGACAGGCACCTGGAAAAAGGCTTTGAGTGGATTGGAGAAATTCAT	3006	Qy	4026	GAAACAGA-GAGACAGCAAGATATGGGCCAAAACAGGATATCTGTGTGTAAGCAGTTCTTCC	4084
				Db	3328	GAAACAGTGAACAGCTGAATATATGGGCCAAAACAGGATATCTGTGTGTAAGCAGTTCTTCC	3387

QY 4085 CC-GCTCAGGGCCAGAGACAGTTGGAAACAGAGGAATATATGGGCCAAACAGGATATCTGTGG 4143  
DB |||||  
QY 3388 CCGGCTCAGGGCCAAAGAACAGATGGAACAGCTGAAATATGGGCCAAACAGGATATCTGTGG 3447  
DB |||||  
QY 4144 TAAGCAGTTCTGTCCCGGCTCAGGGCCAAAGAACAGATGTTCCAGATGCGGTCCCGCC 4203  
DB |||||  
QY 3448 TAAGCAGTTCTGTCCCGGCTCAGGGCCAAAGAACAGATGTTCCAGATGCGGTCCAGCC 3507  
DB |||||  
QY 4204 CTCAGCAGTTCTAGAGAACCATCAGATGTTTCCAGGGTGGCCCAAGGACCTGAAATGAC 4263  
DB |||||  
QY 3508 CTCAGCAGTTCTAGAGAACCATCAGATGTTTCCAGGGTGGCCCAAGGACCTGAAATGAC 3567  
DB |||||  
QY 4264 CTTGTCCCTTATTTGAACATAACCATCAGTTGCGCTTCTCGCTTCTGTTCGGCGGCTTCTG 4323  
DB |||||  
QY 3568 CTTGTCCCTTATTTGAACATAACCATCAGTTGCGCTTCTCGCTTCTGTTCGGCGGCTTCTG 3627  
DB |||||  
QY 4324 CTCCCGAGCTCAATAAAGAGCCACAAACCCCTCAGTCGGGCGCCAGTCTCCGATAG 4383  
DB |||||  
QY 3628 CTCCCGAGCTCAATAAAGAGCCACAAACCCCTCAGTCGGGCGCCAGTCTCCGATG 3687  
DB |||||  
QY 4384 ACTGCGTCCCGGCTACCCGTTCTCAATAAACCCCTCTTGCACTTGCATCCGACTCGT 4443  
DB |||||  
QY 3688 ACTGAGTCGCGGGTACCCGTTGATCAATAAACCCCTCTTGCACTTGCATCCGACTG 3747  
DB |||||  
QY 4444 GGTCTCGCTGTTCTTGGGAGGGTCT-CTCTGAGTGAATGACTACCCGTCAAGCGGGTCT 4502  
DB |||||  
QY 3748 GGTCTCGCTGTTCTTGGGAGGGTCTCTCTGAGTGAATGACTACCCGTCAAGCGGGTCT 3804  
DB |||||  
QY 4503 TTCAGTTTCTCCACCTACAGGTCTCAATAACCTCTGATGTCGGCGAGGACTCCG 4562  
DB |||||  
QY 3805 ----- 3804  
QY 4563 TCAGCCCGGTTTGTATTAATAAAATGCAAGAACAGTGTTCCTTCAAGCCAGACTAC 4622  
DB |||||  
QY 3805 ----- 3804  
QY 4623 ATCTGACTCTCGGCTTATATAAAGATGTTGAAGGGCTGTGGACTATCTGCCACAG 4682  
DB |||||  
QY 3805 ----- 3804  
QY 4683 ACTTTTAAAGATTTTATGCTCTCTGGATGAGGATTTAGTCAATCTATCTCTGTCTATT 4742  
DB |||||  
QY 3805 ----- 3804  
QY 4743 TTGCTGGCTTCTCGGTATTTTAAATTTCTAGTTTGCACTCTCTTCTGAGAGCAGCGGA 4802  
DB |||||  
QY 3805 ----- 3804  
QY 4803 TTGAGAGTAGTTAATACTCTGAGGGCAGGCTTCTGTGAAAAGGTTGCTGGGCTCAGTG 4862  
DB |||||  
QY 3805 ----- 3804  
QY 4863 TGAGATTTTGCCATAAAGGGTCTGCCCTGTGTACAGACAGATCGGAATCTAGAGT 4922  
DB |||||  
QY 3805 ----- 3804  
QY 4923 GCATACTCAGAGTCCCGCGGTTCCGGGGCTCTGTATCTCAGGGCATCTTTGCCCTAGAGAT 4982  
DB |||||  
QY 3805 ----- 3804  
QY 4983 CCTCTACGCCGAGCGCATCGTGGCGGGTACCGAGCTCGAATTCGTAATCATGTTGTCATAG 5042  
DB |||||  
QY 3805 ----- 3804  
QY 5043 CTGTTTCTGTGAAATTTGTTATCCGCTCAATTCACAACATACGAGCGGAGC 5102  
DB |||||  
QY 3805 ----- 3804  
QY 5103 ATAAAGTGAAGCTGGGGTGCCTAATGAGTGAGCTAACTCACAATTAAATGCGTGGC 5162  
DB |||||  
QY 3805 ----- 3804

QY 5163 TCACCTGCCCGCTTTCAGTTCGGGAAAACCTGTGCTGCCAGCTGCATTAATGAATCGGCCAA 5222  
DB |||||  
QY 3805 ----- 3804  
QY 5223 CGCGCGGGAGAGCGGTTTGGGGCTCTTCCGCTTCTCCGCTTCTCCTCCTCAGCTACGTCG 5282  
DB |||||  
QY 3836 CGCGCGGGAGAGCGGTTTGGGGCTCTTCCGCTTCTCCTCCTCAGCTACGTCG 3895  
DB |||||  
QY 5283 CTGCGCTCGGTCTGCTCGGCTCGGCGAGCGGTATCAGCTCACTCAAAAGCGGTAAATAGCG 5342  
DB |||||  
QY 3896 CTGCGCTCGGTCTGCTCGGCTCGGCGAGCGGTATCAGCTCACTCAAAAGCGGTAAATAGCG 3955  
DB |||||  
QY 5343 TTATCCACAGAAATCAGGGGATAACGCGAGGAAGAACATGTGAGCAAAAGGCCAGCAAAAG 5402  
DB |||||  
QY 3956 TTATCCACAGAAATCAGGGGATAACGCGAGGAAGAACATGTGAGCAAAAGGCCAGCAAAAG 4015  
DB |||||  
QY 5403 GCCAGGAACCGTAAAAAGGCGGCTGCTGCGGTTTTTCCATAGGCTCCGCCCTCCCTGAC 5462  
DB |||||  
QY 4016 GCCAGGAACCGTAAAAAGGCGGCTGCTGCGGTTTTTCCATAGGCTCCGCCCTCCCTGAC 4075  
DB |||||  
QY 5463 GAGCATCAAAAAATCGACCGCTCAAGTCAGAGGTGGCGAAAACCCGACAGGACTATAAAGA 5522  
DB |||||  
QY 4076 GAGCATCAAAAAATCGACCGCTCAAGTCAGAGGTGGCGAAAACCCGACAGGACTATAAAGA 4135  
DB |||||  
QY 5523 TACCAGCGGTTTCCCTCGGAAAGCTCCCTCGTGGCTCTCCTGTTCCGACCTGCGGCTT 5582  
DB |||||  
QY 4136 TACCAGCGGTTTCCCTCGGAAAGCTCCCTCGTGGCTCTCCTGTTCCGACCTGCGGCTT 4195  
DB |||||  
QY 5583 ACCGGATACCTGTCGGCTTCTCCCTCGGGAAGCGTGGCGCTTCTCATAGCTCAGC 5642  
DB |||||  
QY 4196 ACCGGATACCTGTCGGCTTCTCCCTCGGGAAGCGTGGCGCTTCTCATAGCTCAGC 4255  
DB |||||  
QY 5643 TGTAAGTATCTCAGTTCGGTGTAGTGTCTCCCTCAAGCTGGGCTGTGTGCAAGAACCC 5702  
DB |||||  
QY 4256 TGTAAGTATCTCAGTTCGGTGTAGTGTCTCCCTCAAGCTGGGCTGTGTGCAAGAACCC 4315  
DB |||||  
QY 5703 CCGCTTCAGCGCCAGCCGCTTATCGGTTAACTATCTCTGTGAGTCAAAACCCGTA 5762  
DB |||||  
QY 4316 CCGCTTCAGCGCCAGCCGCTTATCGGTTAACTATCTCTGTGAGTCAAAACCCGTA 4375  
DB |||||  
QY 5763 AGACACGACTTATCGCCACTGGGACGACCCACTGGTAACAGGATTTAGCAGAGCGAGTAT 5822  
DB |||||  
QY 4376 AGACACGACTTATCGCCACTGGGACGACCCACTGGTAACAGGATTTAGCAGAGCGAGTAT 4435  
DB |||||  
QY 5823 GTAGCGGTGTCTCAGAGTCTTCAAGTGTGGCTTAACCTACCGCTACCTAGAGGACA 5882  
DB |||||  
QY 4436 GTAGCGGTGTCTCAGAGTCTTCAAGTGTGGCTTAACCTACCGCTACCTAGAGGACA 4495  
DB |||||  
QY 5883 GTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAGTTGGTAGCTCT 5942  
DB |||||  
QY 4496 GTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAGTTGGTAGCTCT 4555  
DB |||||  
QY 5943 TGATCCGGCAAAACAAACCAACCGCTGGTAGCGGTGGTTTTTTTGGTTGCAAGCAGCAGATT 6002  
DB |||||  
QY 4556 TGATCCGGCAAAACAAACCAACCGCTGGTAGCGGTGGTTTTTTTGGTTGCAAGCAGCAGATT 4615  
DB |||||  
QY 6003 ACGGCAGAAAAAAGGATCTCAAGAGAGTCTTGTGATCTTTCTACGGGCTCTGACGCT 6062  
DB |||||  
QY 4616 ACGGCAGAAAAAAGGATCTCAAGAGAGTCTTGTGATCTTTCTACGGGCTCTGACGCT 4675  
DB |||||  
QY 6063 CAGTGGAAACGAAACTCACCGTTAAGGATTTGGTCTAGAGATTATCAAAAGGATCTTC 6122  
DB |||||  
QY 4676 CAGTGGAAACGAAACTCACCGTTAAGGATTTGGTCTAGAGATTATCAAAAGGATCTTC 4735  
DB |||||  
QY 6123 ACCTAGATCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAA 6182  
DB |||||  
QY 4736 ACCTAGATCTTTT- - - - - GCGCGCGCGCGCAAAATCAATCTAAAGTATATATAGTAA 4788  
DB |||||  
QY 6183 ACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGACCACTATCTCAGCGATCTGTCTA 6242  
DB |||||  
QY 4789 ACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGACCACTATCTCAGCGATCTGTCTA 4848  
DB |||||  
QY 6243 TTTGCTTTCATCCATAGTTGCTGACTCCCGCTCGTGTAGATAACTAGCAGGAGGCG 6302

Db 4849 TTTTCGTTTCCATAGTTGCTGACTCCCGCTGCTAGATACATGATAGCGAGGCG 4908  
 QY 6303 TTACCATCTGGCCCGAGTGTGCAATGATACCGGAGACCCAGCTCACGGCTCCAGAT 6362  
 Db 4909 TTACCATCTGGCCCGAGTGTGCAATGATACCGGAGACCCAGCTCACGGCTCCAGAT 4968  
 QY 6363 TTATCAGCAATAAACCAGCAGCAGCGGAAGGCGGAGAGTGGTCTGCAACTTTA 6422  
 Db 4969 TTATCAGCAATAAACCAGCAGCAGCGGAAGGCGGAGAGTGGTCTGCAACTTTA 5028  
 QY 6423 TCGGCTTCCATCAGTCTAATTAATGTTGCGGGAAGCTAGAGTAAGTTCGCGAGTT 6482  
 Db 5029 TCGGCTTCCATCAGTCTAATTAATGTTGCGGGAAGCTAGAGTAAGTTCGCGAGTT 5088  
 QY 6483 AATAGTTTGGCAACCTTGTGTCATGCTACAGGC-TGTTGGTGTACGCTCGTCTTT 6541  
 Db 5089 AATAGTTTGGCAACCTTGTGTCATGCTACAGGCATCGTGTGTACGCTCGTCTTT 5148  
 QY 6542 GGTATGGCTTCATTTCAGCTCCGCTTCCCAACGATCAAGCGAGTTACATGATCCCGCATG 6601  
 Db 5149 GGTATGGCTTCATTTCAGCTCCGCTTCCCAACGATCAAGCGAGTTACATGATCCCGCATG 5208  
 QY 6602 TTTGTCAAAAAGCGGTTAGCTCTCTTCGCTCCGATCGTTGTTCAGAAAGTAAAGTTGGCC 6661  
 Db 5209 TTTGTCAAAAAGCGGTTAGCTCTCTTCGCTCCGATCGTTGTTCAGAAAGTAAAGTTGGCC 5268  
 QY 6662 GCAGTGTATCATCTATGTTATGGCAGCACTGATCAATAATTCCTTACTGTCATGCCATCC 6721  
 Db 5269 GCAGTGTATCATCTATGTTATGGCAGCACTGATCAATAATTCCTTACTGTCATGCCATCC 5328  
 QY 6722 GTAAGATGCTTTTCTGTGAGTGTAGTACTCAACCAAGTCAATCTGAGAATAGTGTATG 6781  
 Db 5329 GTAAGATGCTTTTCTGTGAGTGTAGTACTCAACCAAGTCAATCTGAGAATAGTGTATG 5388  
 QY 6782 CGCGCAGCGAGTTGCTTCTTCGCGGCTCAATAACGGGATAATACCGCGCCACATAGCAGA 6841  
 Db 5389 CGCGCAGCGAGTTGCTTCTTCGCGGCTCAATAACGGGATAATACCGCGCCACATAGCAGA 5448  
 QY 6842 ACTTTAAAGTGTCTATCATTTGAAACGTTCTTCGGGCGGAAAACCTCTCAAGGATCTTA 6901  
 Db 5449 ACTTTAAAGTGTCTATCATTTGAAACGTTCTTCGGGCGGAAAACCTCTCAAGGATCTTA 5508  
 QY 6902 CGCGCTGTGAGTCCAGTTCGATTAACCCACTCGTGACCCCAACTGATCTTCAGCATCT 6961  
 Db 5509 CGCGCTGTGAGTCCAGTTCGATTAACCCACTCGTGACCCCAACTGATCTTCAGCATCT 5568  
 QY 6962 TTTACTTTTACCAGCGCTTCTTCGGGTGAGCAAAAACAGGAAGCAAAATCCCGCAAAAAG 7021  
 Db 5569 TTTACTTTTACCAGCGCTTCTTCGGGTGAGCAAAAACAGGAAGCAAAATCCCGCAAAAAG 5628  
 QY 7022 GGAATAAGCGGACACGGAATTTGAAATCTATCTCTTCTTTTCAATATTATTCGA 7081  
 Db 5629 GGAATAAGCGGACACGGAATTTGAAATCTATCTCTTCTTTTCAATATTATTCGA 5688  
 QY 7082 AGCATTTATCAGGGTTATTTGTTCTCATGCGGATACATATTTGAATGTTATTTAGAAAAAT 7141  
 Db 5689 AGCATTTATCAGGGTTATTTGTTCTCATGCGGATACATATTTGAATGTTATTTAGAAAAAT 5748  
 QY 7142 AAACAAATAGGGGTTCCGCGCACATTTCC 7170  
 Db 5749 AAACAAATAGGGGTTCCGCGCACATTTCC 5777

RESULT 8

ID AAX33182 standard; DNA; 7372 BP.

XX AAX33182;

AC AAX33182;

XX 25-JUN-1999 (first entry)

XX Base sequence of the plasmid pRx-Bcl-xl-bst.

XX Cowpox virus; bsr; viral vector; expression; apoptosis; resistance; crmA;  
 KW bcl-2; bcl-xl; Flp; survivin; IAP; ILP; adenovirus; cancer;  
 KW autoimmune disease; graft rejection reaction; inflammation;  
 KW inflammatory disease; ss.  
 OS Synthetic.  
 OS Homo sapiens.  
 PN WO9913073-A2.  
 XX 18-MAR-1999.  
 XX 07-SEP-1998; 98WO-JP004010.  
 XX 08-SEP-1997; 97JP-00259235.  
 XX (RPRG-) RPR GENCELL ASIA PACIFIC INC.  
 XX Hamada H;  
 XX WPI; 1999-243728/20.  
 XX New apoptosis-resistant virus-sensitive cell.  
 XX Example 2; Page 41-45; 51pp; English.  
 CC The present invention describes an apoptosis-resistant virus-sensitive  
 cell line into which an apoptosis resistance gene has been introduced.  
 CC The recombinant viruses generated are capable of expressing apoptosis-  
 CC associated genes. These can then be used in a variety of diseases for  
 CC which the induction of apoptosis by gene transfer, or where the  
 CC inhibition of harmful apoptosis, is therapeutic. The recombinant viruses  
 CC are useful as vectors for gene therapy which can be applied to cancer  
 CC therapy for destroying cancer cells selectively, the treatment of  
 CC autoimmune diseases and graft rejection reaction, and apoptosis induction  
 CC therapy for inflammatory cells in inflammatory diseases. Prior arts have  
 CC encountered the problem where if an adenovirus vector capable of  
 CC expressing an apoptosis-associated gene is introduced into animal cells,  
 CC the cells producing the virus will be destroyed because the period of  
 CC time required to induce cell death by apoptosis is shorter than that  
 CC required to replicate and produce the virus, resulting in failure to  
 CC obtain a recombinant virus having the integrated apoptosis-associated  
 CC gene. In this invention an apoptosis-resistant 293 cell line (having an  
 CC apoptosis resistant gene introduced) is established and overcomes the  
 CC problem. The present sequence represents the base sequence of the plasmid  
 CC pRx-Bcl-xl-bst, which contains the human Bcl-xl gene, and is used in an  
 CC example from the present invention  
 XX  
 SQ Sequence 7372 BP; 2353 A; 1749 C; 1649 G; 1621 T; 0 U; 0 Other;  
 Query Match 41.6%; Score 3183.2; DB 2; Length 7372;  
 Best Local Similarity 70.7%; Pred. No. 0;  
 Matches 4893; Conservative 0; Mismatches 1503; Indels 528; Gaps 28;  
 QY 843 TTCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCCAGAGACCTGAATGACCTGTGCT 902  
 Db 467 TGCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCCAGAGACCTGAATGACCTGTGCT 526  
 QY 903 TATTTGAATTAACCAATCAGTTCGCTTCTCGCTTCTGTTGCGCGCTTCTGCTCCCGAG 962  
 Db 527 TATTTGAATTAACCAATCAGTTCGCTTCTCGCTTCTGTTGCGCGCTTCTGCTCCCGAG 586  
 QY 963 CTCATAAAGAGCCCAACACCCCTCACTCGCGCGCCAGTCTCCGATAGACTCGGTGCG 1022  
 Db 587 CTCATAAAGAGCCCAACACCCCTCACTCGCGCGCCAGTCTCCGATAGACTCGGTGCG 646  
 QY 1023 CCGGGTACCGGTATTTCCCAATTAAGCCTTCTGCTTGTGATCCGAATCGTGGACTCGC 1082  
 Db 647 CCGGGTACCGGTATTTCCCAATTAAGCCTTCTGCTTGTGATCCGAATCGTGGACTCGC 705  
 QY 1083 TGATCTTTGGGAGGTCTCTCAGATTGATGACTGCTCC-ACCTCGGGGTCTTTCATTT 1141



706	TGTTCCCTTGGGAGGGTCTCTCTCGATGTGATTTGACTACCGCTGACGGGGGTCTTTTCAATTT	765
1142	GGAGGTTCCACCGAGATTTGGAGACCCCTGCCACGAGGACACCGACACCCCCCGCCGGAG	1201
766	GGGGCTCGTCGGGATCGGGAGACCCCTGCCCCAGGACCAACCGA-CCACACCAACCGGAG	824
1202	GTAAGCTGGCCAGCAACTTATCTGTGTCTGTCCGATTGTCTAGTGTCTATGACTGATTTT	1261
825	GTAAGCTGGCCAGCAACTTATCTGTGTCTGTCCGATTGTCTAGTGTCTATGACTGATTTT	884
1262	ATGGCCCTGCGTCCGTACTAGTTAGCTAACTTAGCTCTGTATCTTGGCGGACCCGTGGTGA	1321
885	ATGGCCCTGCGTCCGTACTAGTTAGCTAACTTAGCTCTGTATCTTGGCGGACCCGTGGTGA	944
1322	ACTGACAGATTCCGAAACACCCGGCCGCAACCCCTGGGAGACGTCCACGGGACTTTCGGGGC	1381
945	ACTGACAGATTCCGAAACACCCGGCCGCAACCCCTGGGAGACGTCCACGGGACTTTCGGGGC	1004
1382	CGTTTTTGTGGCCCGACCTGAGTCTCTAAAATCCCGATCGTTTATAGGACTCTTTGGTGCAC	1441
1005	CGTTTTTGTGGCCCGACCTGAGTCTCTAAAATCCCGATCGTTTATAGGACTCTTTGGTGCAC	1064
1442	CCCCTTAGAGGAGGATATGTGGTTCTGCTAGGAGACGAGAACTCTAAAACAGTTTCCGCC	1501
1065	CCCCTTAGAGGAGGATATGTGGTTCTGCTAGGAGACGAGAACTCTAAAACAGTTTCCGCC	1124
1502	TCCGCTCGAATTTTGTCTTTCCGTTTGGGACCGAAGCCGGCCGCGCTCTTGTCTGCTG	1561
1125	TCCGCTCGAATTTTGTCTTTCCGTTTGGGACCGAAGCCGGCCGCGCTCTTGTCTGCTG	1184
1562	CAGCATCGTCTGTGTGTCTCTGTCTGACTGTGTCTGTATTTGTCTGAAATATGCG	1621
1185	CAGCATCGTCTGTGTGTCTCTGTCTGACTGTGTCTGTATTTGTCTGAAATATGCGG	1244
1622	CCCGGGCTAGACTGTTTACCACTCCCTTAAAGTTTACCTTAGCTAGCTACCTGGAAGATGTCGA	1681
1245	CC-----AGACTGTTACCACTCCCTTAAAGTTTACCTTAGCTAGCTACCTGGAAGATGTCGA	1298
1682	GCGGATCGCTCACAAACAGTCGGTAGATGTCAAGAGAGACGTTGGGTTACCTTCTGCTC	1741
1299	GCGGATCGCTCACAAACAGTCGGTAGATGTCAAGAGAGACGTTGGGTTACCTTCTGCTC	1358
1742	TGCAGAATGGCAACCTTTTAACTGTCGATGGCCGCGAGACGGCACTTTTAAACCGAGACCT	1801
1359	TGCAGAATGGCAACCTTTTAACTGTCGATGGCCGCGAGACGGCACTTTTAAACCGAGACCT	1418
1802	CATCACCCAGGTTAAGATCAAGTCTTTTACCTGGCCCGCAGTGGACACCCAGACCAAGT	1861
1419	CATCACCCAGGTTAAGATCAAGTCTTTTACCTGGCCCGCAGTGGACACCCAGACCAAGT	1478
1862	CCCCTACATCGTAGCTGGGAAGCTTTGGCTTTTGAACCCCTCCCTGGGTCAAGCCCTT	1921
1479	CCCCTACATCGTAGCTGGGAAGCTTTGGCTTTTGAACCCCTCCCTGGGTCAAGCCCTT	1538
1922	TGTAACCCCTAAGCCTCCGCTCTCTTCTCCATCGCCCGCTCTCTCCCTTGAAC	1981
1539	TGTAACCCCTAAGCCTCCGCTCTCTTCTCCATCGCCCGCTCTCTCCCTTGAAC	1598
1982	TCCTCGTTGACCCCGCTCGATCTCCCTTTATCGAGCCCTCACTCTCTCTTAGGCGC	2041
1599	TCCTCGTTGACCCCGCTCGATCTCCCTTTATCGAGCCCTCACTCTCTCTTAGGCGC	1658
2042	CCCCATATGGCCATATGAGATCTTATATGGGACACCCCGCCCTTCTAACTTCCCTGA	2101
1659	-----CCATATGAGATCTTATATGGGGACACCCCGCCCTTCTAACTTCCCTGA	1708
2102	CCCTGCATGACAAGAGTTACTAACAGCCCTCTCTCCAAGCTCACCTTACAGGCT-TCCTA	2160
1709	CCCTGCATGACAAGAGTTACTAACAGCCCTCTCTCCAAGCTCACCTTACAGGCTCTCTA	1768
2161	CTTAGTCCAGCAAGAGTCTGGAGACCTTGGCGGACGCTTACCAAGAACACTGAGACCG	2220
1769	CTTAGTCCAGCAAGAGTCTGGAGACCTTGGCGGACGCTTACCAAGAACACTGAGACCG	1828

Qy	2221	ACCGGTGGTACTCTCA	CCCTTTACCGAGTCGGCGACACAGTGTGGTCCGCCGACACAGAC	2228
Db	1829	ACCGGTGGTACCTCA	CCCTTTACCGAGTCGGCGACACAGTGTGGTCCGCCGACACAGAC	1888
Qy	2281	TAGAAGCTAGAACCT	CGCTCGTGGAAAGACCTTACACAGTCTCTGTACCAACCCCAACCGC	2340
Db	1889	TAGAAGCTAGAACCT	CGCTCGTGGAAAGACCTTACACAGTCTCTGTACCAACCCCAACCGC	1948
Qy	2341	CCTCAAGTAGACGG	CATCGACCTTTGGATACACGCGCCGCCACGTTGAAGCTGCCGACCC	2400
Db	1949	CCTCAAGTAGACGG	CATCGACCTTTGGATACACGCGCCGCCACGTTGAAGCTGCCGACCC	2008
Qy	2401	CGGGGGTGGACCA	TCTCTCTAGACTGGCATGGGATGAGCTGTATCATCTCTTCTTCTTGGTGA	2460
Db	2009	CGGGGGTGGACCA	TCTCTCTAGACTGGCATGGGATGAGCTGTATCATCTCTTCTTCTTGGTGA	2068
Qy	2461	GCACAGCTACAG	-----	2473
Db	2069	AAAAATGTCTCAG	AGCAACCGGAGCTGGTGGTGTGACTTTCTCTCTTCAAGCTTTTCCCA	2128
Qy	2474	-----	-----	2473
Db	2129	GAAGGATACAGCT	GGAGTCAGTTTAGTGATGTGGAAAGACACAGGACTCAGGCCCCGACA	2188
Qy	2474	-----	-----	2480
Db	2189	AGGACTGAATCG	AGATGGAGACCCCACTGCATCAATGGCAACCCATCTCTGGCACT	2248
Qy	2481	CTCGGACATCCAG	TGACCCAGAGCCCAAGCAGCTTGAGCGCCAGCGTGGGTGACACAGT	2540
Db	2249	GGCAGACAGCCC	CGCGGTGAATGGAGCACTGCGCACAGCAGCAGTTTGGATGCCCGGA	2308
Qy	2541	GACCATCACCTGT	AAAGGC-----	2577
Db	2309	GGTGATCCCAT	GGCAGCAGTAAAGCAGCGCTGAGGGAGGCAAGCGACGAGTTTGAAT	2368
Qy	2578	TCTGTAGCTTGG	TACACGACGACGAGT-----	2634
Db	2369	CGGTACCGCGCG	CATTACGTACCTGACATCCAGCTCCACATCACCCACGGGACAGC	2428
Qy	2635	ACATCCACCGGCA	CACTGGTGTGCCAGACAGATTACGGGTAGCGGTA-----	2687
Db	2429	ATATCAGAGCTT	TTCACACAGGTAGTGAATGAATCTTCTCCGGGATGGGTAACTGGGGTCG	2488
Qy	2688	TACCGACTTACCT	TCCACATCAGACGCTCCAGCCAGCAGGACATCGCCACCTACTACTG	2747
Db	2489	CATTGTGGCC	TTTTTCTCTTCGCGGGGCACTGTGCGTGGAAAGCGTAGACAAGAGAT	2548
Qy	2748	CCAGCAATA-----	-----	2756
Db	2549	GCAGGTATTGGT	GAGTCGGATCGCAGCTTGGATGGCCACTTACTGTAATGACCACTAGA	2608
Qy	2757	---TAGCCTCTAT	CGGTCTGGCCCAAGGGAACAAAGGTGGAATCAACAGAGGTGGCTC	2813
Db	2609	GCCTTGGATCC	AGGAGAACGGCGCTGGGATCTTTTGTGAACTCTATGGGAAACAATGC	2668
Qy	2814	AGGATC---GGT	GTGATCGGCTCTGGTGGCTCAGGATCGGAGGTCCTCACTGGTGGAGAG	2870
Db	2669	AGCAGCCGAGAG	CGGCAAGGGCCAGGAACGCTTCAACCGCTGTTCTGTACGGGCATGAC	2728
Qy	2871	CGGTGGAGGTGT	GTGCAACTCGCGGTCTCTCGCCTG-----	2910
Db	2729	TGTGGCCGGTGT	CTCTGTGGCTCACTCTTCAGTCGGAATGAAAAAAGAGAT	2788
Qy	2911	TCCTGTCCGCAT	CTGGCTTCGATTTTCCACACATATTTGGATGAGTTGGGTGAGACAGGCA	2970
Db	2789	TCGAGCTCGGT	ACCGGGCGCCGGATCTTTAAATAGCGCGCGGTACAATTTGT	2848
Qy	2971	CCTGGAAAAAGT	CTTGATGGATTGGAGAAAAATCATCCAGATACGATGACATTAATCTAT	3030
Db	2849	TAACTACTAACT	AAAGTACGACAGTTCAGCGTTTCCCTCTAGCGGATCAATTTCCCTCTCCCT	2908





QY 7173 GAAAGTGCACCTGACGCTCTAAGAAACCAATTATTATCATGACATTAACTATATAAATA 7232  
 |||||  
 Db 7289 GAAAGTGCACCTGACGCTCTAAGAAACCAATTATTATCATGACATTAACTATATAAATA 7348  
 |||||  
 QY 7233 GCGTATCACGAGGCCCTTTCGTC 7256  
 |||||  
 Db 7349 GCGTATCACGAGGCCCTTTCGTC 7372  
 |||||

RESULT 9  
 AAL57215  
 ID AAL57215 standard; DNA; 6253 BP.  
 XX  
 AC AAL57215;  
 XX  
 DT 25-SEP-2003 (first entry)  
 XX  
 DE MIG retroviral vector DNA sequence.  
 XX  
 KW MIG retroviral vector; lymphocyte; cytostatic; antiallergic;  
 KW immunosuppressive; antimicrobial; anti-HIV; virucidal; gene therapy;  
 KW cancer; allergy; autoimmune disorder; infectious disease; HIV; AIDS;  
 KW OTII T cell receptor; TCR; ds.  
 XX  
 OS Unidentified.  
 XX  
 PN WO2003050262-A2.  
 XX  
 PD 19-JUN-2003.  
 XX  
 PF 10-DEC-2002; 2002WO-US039527.  
 XX  
 PR 10-DEC-2001; 2001US-033937SP.  
 PR 08-JUL-2002; 2002US-0394803P.  
 XX  
 PA (CALY ) CALIFORNIA INST OF TECHNOLOGY.  
 XX  
 PI Yang L, Van Parijs L, Baltimore D;  
 XX  
 DR WPI; 2003-532909/50.  
 XX  
 PT Generating a lymphocyte with a unique antigen specificity in a mammal  
 PT utilizing an antigen-specific polynucleotide, useful for diagnosing and  
 PT treating cancer, allergy, autoimmune disorders and infectious disease  
 PT such as HIV/AIDS.  
 XX  
 PS Disclosure; Fig 5; 59pp; English.  
 XX  
 CC This invention relates to the generation of a lymphocyte with a unique  
 CC antigen specificity in a mammal. The method comprises contacting a  
 CC mammalian stem cell with a polynucleotide delivery system comprising an  
 CC antigen specific polynucleotide and transferring the mammalian stem cell  
 CC into the mammal. The antigen specific polynucleotide encodes an antigen  
 CC specific polypeptide. The method of the invention may be used to produce  
 CC cytostatic, antiallergic, immunosuppressive, antimicrobial, anti-HIV and  
 CC virucidal activities via gene therapy. The methods and compositions of  
 CC the present invention may be useful for diagnosing and treating cancer,  
 CC allergy, autoimmune disorders and infectious diseases, such as HIV/AIDS.  
 CC The present sequence is that of the MIG retroviral expression vector used  
 CC to express both green fluorescent protein (GFP), to mark infected cells,  
 CC and a gene of interest. This was used to express the gene for the OTII T  
 CC cell receptor (TCR) during an in vitro demonstration of functional  
 CC expression of antigen specific TCRs in example 1 of the specification  
 XX  
 SQ Sequence 6253 BP; 1469 A; 1750 C; 1611 G; 1423 T; 0 U; 0 Other;  
 Query Match 41.0%; Score 3139.6; DB 9; Length 6253;  
 Best Local Similarity 71.3%; Pred. No. 0;  
 Matches 50%; Conservative 0; Mismatches 889; Indels 1153; Gaps 25;  
 QY 552 TGAAGACCCACCTCTAGCTTTGGCAAGCTAGCTTTAAGTAAGCCATTGCAAGGCAT 611  
 |||||  
 Db 1 TGAAGACCCACCTCTAGCTTTGGCAAGCTAGCTTTAAGTAAGCCATTGCAAGGCAT 60  
 |||||

QY 612 GGAAATACATACTAGAAATAGAAAGTTTCAGATCAAGGTTAGGAAACAGAGAGACAGCA 671  
 |||||  
 Db 61 GGAAATACATACTAGAAATAGAAAGTTTCAGATCAAGGTTAGGAAACAGAGAGACAGC- 119  
 |||||  
 QY 672 GAAATATGGCCCAACACAGGATATCTGTGTTAAGCAGTTCTGTCCCGCTCAGGCCCAAGAA 731  
 |||||  
 Db 120 ----- 119  
 QY 732 CAGTTGGAAACAGAGAGAAATATGGGCCCAACACAGGATATCTGTGTTAAGCAGTTCTGTCCCGC 791  
 |||||  
 Db 120 -----AGAAATATGGGCCCAACACAGGATATCTGTGTTAAGCAGTTCTGTCCCGC 166  
 |||||  
 QY 792 GCTCAGGGCCAAAGAACAGATGTCCTCCAGATCGCGTCCCGCCTCAGCAGTTTCTAGAGA 851  
 |||||  
 Db 167 GCTCAGGGCCAAAGAACAGATGTCCTCCAGATCGCGTCCCGCCTCAGCAGTTTCTAGAGA 226  
 |||||  
 QY 852 ACCATCAGATGTTTCCAGGGTGCCTCCCAAGGACCTG- AAATGACCTGTGCTTATTTGAA 910  
 |||||  
 Db 227 ACCATCAGATGTTTCCAGGGTGCCTCCCAAGGACCTGAAATGACCTGTGCTTATTTGAA 286  
 |||||  
 QY 911 CTAAACAATCAGTTGCTTCTGCTTCTGTTGCGGGCTTCTGCTCCCGAGCTCAATAA 970  
 |||||  
 Db 287 CTAAACAATCAGTTGCTTCTGCTTCTGTTGCGGGCTTCTGCTCCCGAGCTCAATAA 346  
 |||||  
 QY 971 AAGAGCCACAAACCCCTCACTCGGCGCGCAGTCTCTCGATAGACTGCGTCCCGGGTA 1030  
 |||||  
 Db 347 AAGAGCCACAAACCCCTCACTCGGCGCGCAGTCTCTCGATAGACTGCGTCCCGCC-GGTA 405  
 |||||  
 QY 1031 CCCGTATTCCCAATAAAGCCTTCTGCTTGTTCATCCGAATCGTGGACTCGTGATTCCTT 1090  
 |||||  
 Db 406 CCCGTATTCCCAATAAAGCCTTCTGCTTGTTCATCCGAATCGTGGACTCGTGATTCCTT 465  
 |||||  
 QY 1091 GGGAGGGTCTCTCAGATTGATTGATGCTGCGGAGCTGCGGGTCTTTCATTTGGAGTTCC 1150  
 |||||  
 Db 466 GGGAGGGTCTCTCAGATTGATTGATGCTGCGGAGCTGCGGGTCTTTCATTTGGAGTTCC 525  
 |||||  
 QY 1151 ACCGAGATTTGGAGACCCCTGCGGAGGACCAACCGGACCCCGCGGGAGGTAGACTGG 1210  
 |||||  
 Db 526 ACCGAGATTTGGAGACCCCTGCGGAGGACCAACCGGACCCCGCGGGAGGTAGACTGG 585  
 |||||  
 QY 1211 CCAGCAACTTATCTGTGCTGTC-----CGATTGTCTAGTGTCTATGACTGATTTTA 1262  
 |||||  
 Db 586 CCAGCGGTGTTTGTGCTGCTCTGCTGTTGCGTGTGTTGTCGGGCATCTAATGTT 645  
 |||||  
 QY 1263 TCGGCTGCGTGGTACTAGTTAGTAACTAGTCTGATCTGATCTGCGGAGACCTGTGGAA 1322  
 |||||  
 Db 646 TCGGCTGCGTGGTACTAGTTAGTAACTAGTCTGATCTGATCTGCGGAGACCTGTGGAA 705  
 |||||  
 QY 1323 CTGACGAGTTGCGAAACACCCGCGCCCAACCTCGGAGAGCTGCCAGGAGCTTCGGGGGC 1382  
 |||||  
 Db 706 CTGACGAGTTGCGAAACACCCGCGCCCAACCTCGGAGAGCTGCCAGGAGCTTCGGGGGC 765  
 |||||  
 QY 1383 GTTTTTGTGGCCCGACCTGAGTCTCTAAATCCCGATCGTTTAGGACTCTTTGGTGCACCC 1442  
 |||||  
 Db 766 GTTTTTGTGGCCCGACCTGAGGAGGAGTGTGGAATCCGACCCCGTCA----- 818  
 |||||  
 QY 1443 CCCTTAGAGAGGATATGTTGTTCTGTGAGAGAGAGAACTTAAACAGTTCCCGCT 1502  
 |||||  
 Db 819 -----GGATATGTTGTTCTGTGAGAGAGAGAACTTAAACAGTTCCCGCT 866  
 |||||  
 QY 1503 CCGTCTGAATTTTGTGTTTGGGACCGGACCGCGCGCTCTGTT---CTGC 1559  
 |||||  
 Db 867 CCGTCTGAATTTTGTGTTTGGGACCGGACCGCGCGCTCTGTT---CTGC 926  
 |||||  
 QY 1560 TGCAGCATGTTCTGTGTTGTTCTCTGTCTGACTGTGTTTCTGTATTTGTTGTAATAATG 1619  
 |||||  
 Db 927 TGCAGCATGTTCTGTGTTGTTCTCTGTCTGACTGTGTTTCTGTATTTGTTGTAATAATG 986  
 |||||  
 QY 1620 GSCCGGGCTAGACTGTTTACCACTCCCTTAAGTTTGAACCTTAGGTCACCTGGAAGATGTC 1679  
 |||||  
 Db 987 GGCC-----AGACTGTTTACCACTCCCTTAAGTTTGAACCTTAGGTCACCTGGAAGATGTC 1040  
 |||||

QY 1680 GAGCGATCGCTCAACACAGTCGGTAGATGTCAAGAGAGACGTTGGGTTACCTTTGCG 1739  
DB |||||  
QY 1041 GAGCGATCGCTCAACACAGTCGGTAGATGTCAAGAGAGACGTTGGGTTACCTTTGCG 1100  
DB |||||  
QY 1740 TCTGCAGAAATGSCCAACCTTTAACTCGCATGCGCGAGAGCGGACCTTTAAACCGAGAC 1799  
DB |||||  
QY 1101 TCTGCAGAAATGSCCAACCTTTAACTCGCATGCGCGAGAGCGGACCTTTAAACCGAGAC 1160  
DB |||||  
QY 1800 CTCAFCACCCAGGTTAAGATCAAGGTCTTTTCACTGGCGCCGCATGGACACCCAGACCCAG 1859  
DB |||||  
QY 1161 CTCAFCACCCAGGTTAAGATCAAGGTCTTTTCACTGGCGCCGCATGGACACCCAGACCCAG 1220  
DB |||||  
QY 1860 GTCCCTCATAGTCGACCTGGAAAGCTTTGGCTTTTGAACCCCTTCCCTCGGTCGAAGGCC 1919  
DB |||||  
QY 1221 GTCCCTCATAGTCGACCTGGAAAGCTTTGGCTTTTGAACCCCTTCCCTCGGTCGAAGGCC 1280  
DB |||||  
QY 1920 TTTGTACACCCCTTAAGCCTCGGCTCTCTCTCTCCATCGCGCCGCTCTCTCCCTTGAA 1979  
DB |||||  
QY 1281 TTTGTACACCCCTTAAGCCTCGGCTCTCTCTCTCCATCGCGCCGCTCTCTCCCTTGAA 1340  
DB |||||  
QY 1980 CCTCTCGTTGACCCCGCTCGATCTCTCTCTTTATCCAGCCCTCACTCTCTTAGGC 2039  
DB |||||  
QY 1341 CCTCTCGTTGACCCCGCTCGATCTCTCTCTTTATCCAGCCCTCACTCTCTTAGGC 1400  
DB |||||  
QY 2040 GCGCCCATATGGCCATATAGATCTTATATGGGGCACCCCGCCCTTTGTAAACTTCCCT 2099  
DB |||||  
QY 1401 GCGGAGAT-----CT 1410  
DB |||||  
QY 2100 GACCTGATGATGACAAGAGTTACTTAAAGCCCTCTCTTCAAGCTCACTTACAGGCTTCT 2159  
DB |||||  
QY 1411 CTCGAGGACGTTTAAACGAGTTTAAACGAGCGCGCGCAAGCT----- 1455  
DB |||||  
QY 2160 ACTTAGTCCAGCAGAACTCTGGAGACCTCTGGCGGCGAGCTTACCAAGAACAACTGGAGC 2219  
DB |||||  
QY 1456 -----TGACGAATTCGCGCCCTCTCTCTCTCCCTCCCGCCCTTACTGGCCGAA 1505  
DB |||||  
QY 2220 GACCGGTGGTACTCACCTTACCGAGTCGCGACACAGTGTGGGTGCGCGACACCGA 2279  
DB |||||  
QY 1506 GCGCGTTGGAATTAAGCGCGGTGTGGTTGTCTATATGTATTTTCCACCATAATTGCGGT 1565  
DB |||||  
QY 2280 CTAGAACCTTAGAACCTCGCTGGAAAGACCTTACAGAGTCGTGTGACACCCCGACCG 2339  
DB |||||  
QY 1566 CTTTGTGGCAATGTGAGGCGCGGAAACCTTGGCCCTGTCTCTTGACAGACATTTCTAGGG 1625  
DB |||||  
QY 2340 CCTCAAGATGAGCGCATCGAGCTTGAATACACCGCCCGACGTAAGGCTGCGCGAC 2399  
DB |||||  
QY 1626 GTCTTTCCCTCTCCCAAGAAATGCAAGGTCTGTGAATGTGTGAAGGAAGAGTTTC 1685  
DB |||||  
QY 2400 CCGGGGTGGACCATCTCTAGACTGCCATGGGATGGAGCTGTATCATCTCTCTTCTTGGT 2459  
DB |||||  
QY 1686 CTCTG---GAAGCTTCTTGAAGACAAACAACGTCTGTAGCGACCTTTGC----- 1732  
DB |||||  
QY 2460 AGCAACAGCTACAGTGTCTCACTCCGACATCCAGCTGACCCAGAGCCCAAGAGCCTGAG 2519  
DB |||||  
QY 1733 -----AGGACGCGAAACCCCGACCTGGCGACAGGTGCTCTGCGCGCCAAAGCCACCGT 1787  
DB |||||  
QY 2520 CGCCAGCGTGGTGCAGAGTGACCATCACTGTAAAGCCAGTCAGGATGGGTACTTTC 2579  
DB |||||  
QY 1788 TATTAAGATACACTGTCAAGGGGGGCAAAACCCAGTGCACAGTTGTGAGTTGGATAGTTG 1847  
DB |||||  
QY 2580 TGTAGCTTGGTACAGAGAGAGGAGGTAAGGCTTCAAGGCTGCTACTGTGACATC 2639  
DB |||||  
QY 1848 TGGAAAGAGTCA----- 1859  
DB |||||  
QY 2640 CACCGGCACACTGTGTGCCAAGCAGATTCAGCGGTAGCGGTAGCGGTACCGACTTCAAC 2699  
DB |||||  
QY 1860 ----- 1859  
DB |||||  
QY 2700 CTTTACCATACAGCCTCCAGCCAGAGGACATCGCCACCTACTACTCCAGGCAATATAG 2759  
DB |||||  
QY 1860 -----AATGGCT 1866  
DB |||||  
QY 2760 CCTCTACGCTGTTGCGGCCAAGGAGCAAGGTGGAAATCAACAGAGGTGGCTCAGGATC 2819  
DB |||||

DB |||||  
QY 1867 CTCCTCAAGCGTATTTCAACAAGGGGCTGAAGGATGC-----CCAGAAGGTATCCCACTT 1919  
DB |||||  
QY 2820 GGGTGGATCCGGCTCTGGTGGCTCAGGATCGAGGTCCAACTGGTGGAGAGCGGTGGAGG 2879  
DB |||||  
QY 1920 GTATGGGATCTGATCTGGGGCTC-----GCTGCACATGCTTTTACATGTGTTAGT 1970  
DB |||||  
QY 2880 TGTGTGCAACCTGGCCGGTCCCTGGCGCTGTCTGTCTCGCATCTGGCTTTCGATTTCAAC 2939  
DB |||||  
QY 1971 CGAGTTTAAAAACGCTTAGGC---CCCCCGAACCCACGGGAGCGTGTTCCTTTTGAA 2027  
DB |||||  
QY 2940 CACATATTTGGATGAGTTGGGTGAGACAGGACCTTGGAAAAAGTCTTTGAGTGAATTTGAGA 2999  
DB |||||  
QY 2028 AACACGATGATATATGGCCCAACCAAGGCGAGAGAGCTGTTTCAACCGGGTGGTCCCA 2087  
DB |||||  
QY 3000 AATTATCATCCAGTAGCAGTACGATTAATCTATATGCGCGCTCTCTAAGGATAGATTTACAAT 3059  
DB |||||  
QY 2088 TCCTGGTGGAGCTGG----- 2102  
DB |||||  
QY 3060 ATCGCGAGACAACGCCAAGAACACATTTGTTCTGCAAAATGGACAGCTGAGACCCGAAAGA 3119  
DB |||||  
QY 2103 -----ACGGGAGCGTGAAACGGCCACAAGTTTACGCTGTCCGGGAGGG 2145  
DB |||||  
QY 3120 CACCGGGTCTATTTTGTGCAAGCTTTTACTTGGCTTCCCTTGGTTTGTATTTGGGG 3179  
DB |||||  
QY 2146 CGAGGGCGATGCCACCTACGGCAAGCTGACCTGAAAGTTTCTGCAACACCGGCAAGCT 2205  
DB |||||  
QY 3180 CCAAGGACCCCGGTCACCGTCTCCAGTGTCTAAGCCCAACGAGCGCCAGCGCCGAGC 3239  
DB |||||  
QY 2206 GCGCGTGCCTTGGGCCACCT-----CGTGACC 2233  
DB |||||  
QY 3240 ACCAAACCGCGCGCCACCATCGCGTGCAGCCTCTGCTGCGCCCGAGAGCGGCTCG 3299  
DB |||||  
QY 2234 ACCCTGACCTACGGGTGCACTGTCTT---CAGCGCGTACCCCGACCAATGAGCAGCAG 2291  
DB |||||  
QY 3300 GCCAGCGCGCGGGGCGCAGTGCAACAAGGGGGTGGACTTTCGCTTGGATTCCTCAAACT 3359  
DB |||||  
QY 2292 AC-----TTCTTCAAGTCCGCCATGCGCGAAGG 2319  
DB |||||  
QY 3360 CTGTACTCTGTGTGATGGAATCTTTCATCTATGTTGTCTTCTCATCTGCTGTTTCTCT 3419  
DB |||||  
QY 2320 CTACGTCCAGAGCGCACCATCTTCTTTCAGGACGACGCGCAACTACAAGACCCGCGCGA 2379  
DB |||||  
QY 3420 GAGAGTGAAGTTACGAGGAGCGAGCGACCCCGCGTACCAGCAGGCGCCAGAACAGCT 3479  
DB |||||  
QY 2380 GGTGAAGTTGAGGGCGACACCTTGTGTGAACCGCATCGAGCTGAAGGCGATCGACTTCA 2439  
DB |||||  
QY 3480 CTATAACGAGCTCAATCTAGGACGAAGAGAGAGTACGATGTTTGGCAAGAGACGTTGG 3539  
DB |||||  
QY 2440 GGAGGACGGCAACATCTTGGGGCAACAGCTGGAGTACAACCTACAACAGCCCAACGCTCTA 2499  
DB |||||  
QY 3540 CCGGACCTCGAGATGGGGGAAAGCGAGAGGAAGAACCTTCAAGGAGGCTGTATCA 3599  
DB |||||  
QY 2500 TATCATGCCGCAAGCAGAGAAAGACGGCATCAAGCGCAACTTCAAGATCCGCGCAACAT 2559  
DB |||||  
QY 3600 TGAATCGCAGAAAGATGAAGTGGCGGAGCTTACAGTGAGATTGGATGAAGGCGAGCG 3659  
DB |||||  
QY 2560 CGAGGACGGCAGCTGCAAGTCTGCCGACCACTACAGCAGAACACCCCTTCGGGAGCGG 2619  
DB |||||  
QY 3660 CCGGAGGGGCAAGGGGCAAGTGGCTTTTACAGGGTCTCAGTACAGCCCAACAGGACAC 3719  
DB |||||  
QY 2620 CCGCGTGTCTGCTGCCGACCACTACCTGAGCACCCAGTCCGCTTGGAGCAAGACCC 2679  
DB |||||  
QY 3720 CTACGACCGCTTCAATGACAGGCGCTGCCCTTCTGCTTAACTGACAGCGCGCGCGGATCC 3779  
DB |||||  
QY 2680 CAACGAGAGCGGATCACATGCTGCTGAGTTCGTGACCGCGCGCGGATCACCTCA 2739  
DB |||||  
QY 3780 GGATTAGTCCAAATTTGTTAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTGTACTCA 3839  
DB |||||  
QY 2740 CGGATGGACGAGCTGTACAAAGTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2784  
DB |||||  
QY 3840 CAATATCACAGCTGAAGCCCTATAGATGACGAGCCATAGATAAAATAAAGATTTATTT 3899  
DB |||||





QY 6058 ACGCTCAGTGAACGAAACTCAGCTTAAGGGATTTGGTCATGAGATTATCAAAAGGA 6117  
DB 4383 ACGCTCAGTGAACGAAACTCAGCTTAAGGGATTTGGTCATGAGATTATCAAAAGGA 4442  
QY 6118 TCCTCACCTAGATCCTTTTAAATTAATAAATGAAGTTTAAATCAATCTAAAGTATATATG 6177  
DB 4443 TCCTCACCTAGATCCTTTTAAATTAATAAATGAAGTTTAAATCAATCTAAAGTATATATG 4502  
QY 6178 AGTAAACTTGTGTGACAGTTACCAATGCTTAAATCAGTGAGGACCTATCTCAGCGATCT 6237  
DB 4503 AGTAAACTTGTGTGACAGTTACCAATGCTTAAATCAGTGAGGACCTATCTCAGCGATCT 4562  
QY 6238 GTCTATTTCGTTTCATCATAGTTGCTGACTCCCGTCTGCTAGATTAATCACTAGGATACGGG 6297  
DB 4563 GTCTATTTCGTTTCATCATAGTTGCTGACTCCCGTCTGCTAGATTAATCACTAGGATACGGG 4622  
QY 6298 AGGGCTTACCATCTGCGCCAGTGTGCAATGATACCGCGAGACCCAGCTCACCGGCTC 6357  
DB 4623 AGGGCTTACCATCTGCGCCAGTGTGCAATGATACCGCGAGACCCAGCTCACCGGCTC 4682  
QY 6358 CAGATTTATCAGCAATAAACCAGCAGCCGGAAGGGCCGAGCGCAGAAAGTGGTCTCTGCAA 6417  
DB 4683 CAGATTTATCAGCAATAAACCAGCAGCCGGAAGGGCCGAGCGCAGAAAGTGGTCTCTGCAA 4742  
QY 6418 CTTTATCCGCTCCATCCAGTCTATTAAATTTGTTGCGGGAAAGCTAGAGTAAGTTGCTGC 6477  
DB 4743 CTTTATCCGCTCCATCCAGTCTATTAAATTTGTTGCGGGAAAGCTAGAGTAAGTTGCTGC 4802  
QY 6478 CAGTTAATAGTTTGGCAACGTTGTCACATGCTGTACAGGC-TGCTGGTGTACGCTCGT 6536  
DB 4803 CAGTTAATAGTTTGGCAACGTTGTCACATGCTGTACAGGCATCGTGGTGTACGCTCGT 4862  
QY 6537 CGTTTGTGATGGCTTCATTTCAGCTCCGTTCCCAACGATCAAGCGAGTTACATGATCCC 6596  
DB 4863 CGTTTGTGATGGCTTCATTTCAGCTCCGTTCCCAACGATCAAGCGAGTTACATGATCCC 4922  
QY 6597 CCATGTTGTGCAAAAGCGTTAGTCTCTCGTCTCCGATCGTTGTCAAGAGTAAGT 6656  
DB 4923 CCATGTTGTGCAAAAGCGTTAGTCTCTCGTCTCCGATCGTTGTCAAGAGTAAGT 4982  
QY 6657 TGGCGGAGTGTATCACTCATGTTATGGCAGCAGTGCATAATCTCTTACTCTCATGC 6716  
DB 4983 TGGCGGAGTGTATCACTCATGTTATGGCAGCAGTGCATAATCTCTTACTCTCATGC 5042  
QY 6717 CATCGTGAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCAAGTCACTCTGAGAATAGT 6776  
DB 5043 CATCGTGAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCAAGTCACTCTGAGAATAGT 5102  
QY 6777 GTATGCGGCGACCGAGTTGCTTTGCGCGGCTCAATACGGGATAAATACCGGCCACATA 6836  
DB 5103 GTATGCGGCGACCGAGTTGCTTTGCGCGGCTCAATACGGGATAAATACCGGCCACATA 5162  
QY 6837 GCAGAACTTTTAAAGTGTCTATCATTTGAAACCGTTCTTCGGGGGAAACTCTCAAGGA 6896  
DB 5163 GCAGAACTTTTAAAGTGTCTATCATTTGAAACCGTTCTTCGGGGGAAACTCTCAAGGA 5222  
QY 6897 TCTTACCGCTGTGAGATGCTGATGTAACCCACTCTGTCGCCCAACTGATCTTCAG 6956  
DB 5223 TCTTACCGCTGTGAGATGCTGATGTAACCCACTCTGTCGCCCAACTGATCTTCAG 5282  
QY 6957 CATCTTTTACTTTTACCAAGCGTTTCTGGGTGAGCAAAACAGGAAGGCAAAATGCCGCAA 7016  
DB 5283 CATCTTTTACTTTTACCAAGCGTTTCTGGGTGAGCAAAACAGGAAGGCAAAATGCCGCAA 5342  
QY 7017 AAAAGGGAATAAGGGCGACAGGAATGTTGAATCTACTCTCTCTTTTCAATATT 7076  
DB 5343 AAAAGGGAATAAGGGCGACAGGAATGTTGAATCTACTCTCTCTTTTCAATATT 5402  
QY 7077 ATTGAAGCATTTTATCAGGTTTATTGTTCTCATGAGCGGATACATTTTGAATGTTATTAGA 7136  
DB 5403 ATTGAAGCATTTTATCAGGTTTATTGTTCTCATGAGCGGATACATTTTGAATGTTATTAGA 5462  
QY 7137 AAAATAAAATAAGGGGTTCCGGCGCACATTTTCCCGGAAAGTGCCACTGACGTCTAAG 7196

DB 5463 AAAATAAAATAAGGGGTTCCGGCGCACATTTTCCCGGAAAGTGCCACTGACGTCTAAG 5522  
QY 7197 AAAACCATTTATCATGACATTAACTTATAAAAAATAGCGGTATCACGAGGCGCTTTTCGTC 7256  
DB 5523 AAAACCATTTATCATGACATTAACTTATAAAAAATAGCGGTATCACGAGGCGCTTTTCGTC 5582  
QY 7257 TCGCGGCTTCGCGGTGATGACGGTGAACCTCTTGACACATGACGCTCCCGGAGACGGTCA 7316  
DB 5583 TCGCGGCTTCGCGGTGATGACGGTGAACCTCTTGACACATGACGCTCCCGGAGACGGTCA 5642  
QY 7317 CAGCTTCTCTGTAAGCGGATGCGCGGAGCAGACAGCCCGTCAGGCGCGCTCAGCGGCTG 7376  
DB 5643 CAGCTTCTCTGTAAGCGGATGCGCGGAGCAGACAGCCCGTCAGGCGCGCTCAGCGGCTG 5702  
QY 7377 TTGCGGCTGTTCGCGGCTGCGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGC 7436  
DB 5703 TTGCGGCTGTTCGCGGCTGCGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGC 5762  
QY 7437 ACCATATGCGGTGTAATAATACCGCACAGATGCGTAAGGAGAAATACCGCATCAGGCGCC 7496  
DB 5763 ACCATATGCGGTGTAATAATACCGCACAGATGCGTAAGGAGAAATACCGCATCAGGCGCC 5822  
QY 7497 ATTCCGCATTTCAGGCTGCGCAACTGTTGGGAAGGCGGATCGGTGCGGCGCTCTTCGCTAT 7556  
DB 5823 ATTCCGCATTTCAGGCTGCGCAACTGTTGGGAAGGCGGATCGGTGCGGCGCTCTTCGCTAT 5882  
QY 7557 TACGCCAGCTGGCGAAAGGGGATGCTGCAAGCGCATTAAGTTGGGTAACGCCAGGCT 7616  
DB 5883 TACGCCAGCTGGCGAAAGGGGATGCTGCAAGCGCATTAAGTTGGGTAACGCCAGGCT 5942  
QY 7617 TTTCCCGAGTCAGACGTTGTAATAACGACGCGGCGAGTGC 7654  
DB 5943 TTTCCCGAGTCAGACGTTGTAATAACGACGCGGCGAGTGC 5980

## RESULT 10

ADQ80673  
ID ADQ80673 standard; DNA; 7257 BP.  
XX AC ADQ80673;  
XX DX 21-OCT-2004 (first entry)  
XX DE Bcl-XL expression directing retroviral vector, pPB-bclXL-IRES-hrGFP.  
XX KW Survival; neuron; tyrosine hydroxylase; tyrosine 3-monooxygenase; TH;  
XX KW anti-apoptotic; Bcl-XL; neurological disorder; neuroprotective;  
XX KW neurotrophic; antiparkinsonian; transplantation; drug screening;  
XX KW gene profiling; CNS disorder; neurodegenerative disease;  
XX KW pPB-bclXL-IRES-hrGFP; ds.  
XX OS Unidentified.  
XX XX  
XX PN WO2004062554-A2.  
XX XX 29-JUL-2004.  
XX XX 07-JAN-2004; 2004WO-DK000008.  
XX XX 08-JAN-2003; 2003US-0438719P.  
XX PR 11-APR-2003; 2003DK-0000581.  
XX PR 22-APR-2003; 2003US-0464546P.  
XX XX (NSGE-) NSGENE AS.  
XX XX Martinez-Serrano A, Liste I, Villa A;  
XX XX WPI; 2004-544027/52.

Enhancing the survival of neurons or cells expressing tyrosine hydroxylase (TH) for treating neurodegenerative disorders, comprises contacting neurons or TH expressing cells with Bcl-XL or its functional



QY 1986 CGTTCGACCCCGCTCGATCCTCCTTTATCCAGCCCTCACTCTCTCTAGCGCGCCC 2045  
DB |||||  
QY 1605 CGTTCGACCCCGCTCGATCCTCCTTTATCCAGCCCTCACTCTCTCTAGCGCGCCC 1664  
DB |||||  
QY 2046 ATATGGCCATATGAGATCTTATATGGGGCACCCCGCCCTTGTAAACTTCCCTGACCT 2105  
DB |||||  
QY 1665 ATATGGCCATATGAGATCTTATATGGGGCACCCCGCCCTTGTAAACTTCCCTGACCT 1724  
DB |||||  
QY 2106 GACATGACAGAGTGTACTAAAGCCCTCTCTCCAAAGTCTACTTACAGGCT-TCTACTTA 2164  
DB |||||  
QY 1725 GACATGACAGAGTGTACTAAAGCCCTCTCTCCAAAGTCTACTTACAGGCTCTCTACTTA 1784  
DB |||||  
QY 2165 GTCCAGCAGAGTGTGAGACCTCTGGGGCAGCTTACCAAGAACAACTGGACCGCG 2224  
DB |||||  
QY 1785 GTCCAGCAGAGTGTGAGACCTCTGGGGCAG-CGTACCAAGAACAACTGGACCGCG 1843  
DB |||||  
QY 2225 GTGTACCTCACCTTACCGAGTCGGGCACACAGTGTGGGTCCGCCGACACCACTAAAG 2284  
DB |||||  
QY 1844 GTGTACCTCACCTTACCGAGTCGGGCACACAGTGTGGGTCCGCCGACACCACTAAAG 1903  
DB |||||  
QY 2285 AACCTAGAACCTCGCTGGAAGGACCTTACACAGTCTCTGTAACACACCCCGCCCTC 2344  
DB |||||  
QY 1904 AACCTAGAACCTCGCTGGAAGGACCTTACACAGTCTCTGTAACACACCCCGCCCTC 1963  
DB |||||  
QY 2345 AAAGTAGAGGATCGAGCTTGGATACACGGCCGACAGTGAAGGCTGCGACCCCGGG 2404  
DB |||||  
QY 1964 AAAGTAGAGGATCGAGCTTGGATACACGGCCGACAGTGAAGGCTGCGACCCCGGG 2023  
DB |||||  
QY 2405 GGTGACCATCTCTAGACTGCATGGGATGGAGTGTATCATCTCTTCTTGTAGCAA 2464  
DB |||||  
QY 2024 GGTGACCATCTCTAGACTGCAGTCCGATCGAATTTGGTACCGAGCTCGGATC-CACTAGTAA 2082  
DB |||||  
QY 2465 CAGCTACAGGTGTCCACTCCGACATCCAGCTGACCCAGAGCCCAAGCAGCTGAGCGCA 2524  
DB |||||  
QY 2083 CGGCGCCAGTGTCTGGAGCCACTCTATATATAAATGTCTCAGAGCAACCGGAGC 2142  
DB |||||  
QY 2525 CGGTGGGTGACAGAGTGAAC-ATCACTGTGAAGCCAGTCAAGATGTGGGTACTTCTGTA 2583  
DB |||||  
QY 2143 TGGTGGTGTGACTTCTCTCTACAGCTCTCCAGAAAGGATACAGCTGGAGTCAGTTTA 2202  
DB |||||  
QY 2584 GCTTGGTACCAGCAGAACCGAGTAAAGCTCCAAAGCTGTGATCTACTGTGNCATCCACC 2643  
DB |||||  
QY 2203 GCGATGTGGAAGAGAACAGACTGAAGCCCGAGAGAAACTGAACAGAAAGGGAGACCC 2262  
DB |||||  
QY 2644 CGGCACACTGTGTGCCAGCAGATTCAAGCGGTAGCGGTAGCGGTACCGACTTCACTTC 2703  
DB |||||  
QY 2263 CCAGTGCCATCAATGGCAACCATCTCGCACCTGGCGGATAGCCCCGGGTGAATGGAG 2322  
DB |||||  
QY 2704 ACCATCAGAGCTTCCAGCCAGAGACATCGCCACTACTACTGCGACCAATATAGCCTC 2763  
DB |||||  
QY 2323 CCACTGGCCACAGCAGCAGTTTGGATGGCGCGGAGGTACTCCCATGGCAGCAGTGAAGC 2382  
DB |||||  
QY 2764 TATCGGTGTTGGCAGAGGACCAAGT-----GGNAATCAACAGGTTG 2810  
DB |||||  
QY 2383 AAGCGCTGAGAGAGCTGGCATGAGTTGAACTGCGGTACCGGAGCATCTCAGTGTATC 2442  
DB |||||  
QY 2811 CTCAGGATCGGTGATCCGCTCTGTGGTCTCAGGATCGGAG-GTCCAACTGGTGGAGA 2869  
DB |||||  
QY 2443 TAAATCCAGCTTCATATAACCCAGGACAGTATATCAGAGCTTTGAAACAGGTAGTGA 2502  
DB |||||  
QY 2870 CGGTGGAGGTGTTGTGCAACTGGCCGCTCTGCGCTGTCTCTGCTCCGCATCTGGCT 2929  
DB |||||  
QY 2503 ATGAACCTTTTCGGGATGGGTAAACTGGGGTGGCATTTGGGCTCTCTCTCTTTGGGG 2562  
DB |||||  
QY 2930 TCGATTTCCACATATTTGGATGAGTTGGGTGAGACAGGACCTGGAAAGGCTTGTAGT 2989  
DB |||||  
QY 2563 GGGCACTGTGGGTGAAAGCGTAGAACAGGAGATGCAAGTATTTGGTGTAGTGGGATTCGAA 2622  
DB |||||  
QY 2990 GGAATTGGAGAAATTCATCCAGATAGCATGATTAATATGCGCGCTCTCTAAGGATA 3049  
DB |||||  
QY 2623 G--TTGGATGGCCACTACCTGAATGACCACTTAGAGCTTTGGATCCAGGAGAACGGCGG 2680  
DB |||||

QY 3050 GATTACAAATCCGAGACCAACGCCAAGAACACACTTGTCTCTGCAATGGACACCTGA 3109  
DB |||||  
QY 2681 CTGGAGACACTTTTGTGGATCTCTACGGGAACAATACAGACCAGCGGAAAGGCCA 2740  
DB |||||  
QY 3110 GACCCGAAGACACCGGGGTCTATTTTGTGAAGCTTTACTTCGGCTTCCCTGCTTTG 3169  
DB |||||  
QY 2741 GGAGGTTTCAACCGCTGTTCTTACCGGATGACTGTGGCTGTGTAGTTCTGCTGGG 2800  
DB |||||  
QY 3170 CTTATTTGGGGCCAAAGGACCCCGGTACCGTCTCCAGTGTAAAGCCACACGACGCCAG 3229  
DB |||||  
QY 2801 CTCACTCTCAGTCGAAGTGAACAGATTCAGATATCCATCACACTGGGGCGGGA 2860  
DB |||||  
QY 3230 CGCGGACCAACCAACCGGGGCCA-----CCATCGGCTCGCAGC 3271  
DB |||||  
QY 2861 TTCCGGATCCACTAGTAACGGCCGCGCAGTGTGCGAAATCGCCCTCTCCCTCCCGCC 2920  
DB |||||  
QY 3272 CCCTGTCTCTCGCCCGAGAGCGGTCTCGGCAGCGGGGGGGCGAGTGACACAGAGGG 3331  
DB |||||  
QY 2921 CCTTAAAGTTACTGGCCGAGCGCTTGAATTAAGCCGGTGTGTGTTGTCTATATGTG 2980  
DB |||||  
QY 3332 GGCT-----GGACTTCGCTCTGGATC 3352  
DB |||||  
QY 2981 ATTTCCACCATATTTGCGTCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCT 3040  
DB |||||  
QY 3353 CCAACTCTGCTACTGCTGTGATGGAATCTCTTCTCATCTATGTGTCAATCTCACTGCTC 3412  
DB |||||  
QY 3041 CTTGACGAGCATCTCTAGGGGTCTTTCCCTCTCGCCAAAGGAAATGCAAGGTCTGTTGAA 3100  
DB |||||  
QY 3413 TGTTCCTGAGAGTGAAGTTTACAGAG-----AGCGCAGAGCCCCC 3453  
DB |||||  
QY 3101 TGTGTAAGAGAGAGTCTCTCTGGAAGCTTTTGAAGACAACAACTCTGTATGCGAC 3160  
DB |||||  
QY 3454 CGGTACACGAGGGCCAGAACACAGCTCTATAACAGAGCTCAATCTAGACGACGAGAGGAG 3513  
DB |||||  
QY 3161 CTTTTCGAGGAGCGGAAACCCCGCACAGGTGCTCTGCGCGCAAAAGCCAGC 3220  
DB |||||  
QY 3514 TACGATGTTTGGACACAGAGAGCTGGCCGGGACCTTGAGATGGGGGGAAGCCG----- 3567  
DB |||||  
QY 3221 TGTATAAGATACACTGCAAGGGCGCAACACCCAGTGCACGTTGTGAGTTGATAGT 3280  
DB |||||  
QY 3568 -----AGAGGAGAACCTCAGGAGGCTGTACAACTGACAGAAAGATAAGAT 3620  
DB |||||  
QY 3281 TGTGGAAGAGTCAATGGCTCTCCTCAGCGTAGTCAACAGGGGCTGAAAGGATGCCCA 3340  
DB |||||  
QY 3621 GCGCAGGAGCTTACAG--TGAGATTGGGATGAAGC-----GAGCGCCGAGGGGCAAGGG 3674  
DB |||||  
QY 3341 GAAGTACCCCATTTGTATGGAACTGTATCTGGGGCTCGGTGCACATGCTTTACATGTG 3400  
DB |||||  
QY 3675 GCACGATGGCTTTTACAGGGTCTCAGTACAGC-CACCAAGGACACCTTACGACGCCCTTC 3733  
DB |||||  
QY 3401 TTTAGTCGAGTTTAAAAAGCTTAGGCCCCCGCAACCCAGCGGAGCTGTTTCTCTTTG 3460  
DB |||||  
QY 3734 ACATGCGAGGCTCCCTCGCTAACTCGAGCGGGCGCGGATCCGATTTAGTCCAATT 3793  
DB |||||  
QY 3461 AAAAACAGATGATGAGCTTGCCACAGGAATTCATGTGTGAGCAAGCAGATCCTGAAGAAC 3520  
DB |||||  
QY 3794 TGTTAAAGACAGGATATCAGTGG--TCCAGGCTCTAGTTTGTGACTCAACAATATCACCAG 3851  
DB |||||  
QY 3521 ACCGGCTGCGAGGAGATCATGAGCTTCAAGGTGAACCTGGAGGGGCTGGTGAACACAC 3580  
DB |||||  
QY 3852 CTGAAGCTTATAGATAC-GAGCCATAGATAAAATAAAAGATTTTATTTAGTCTTCAGAA 3910  
DB |||||  
QY 3581 GTGTTCCACATGAGGGCTGGCGCAAGGGCAACATCCTGTTCCGCAACACAGCTGGTGAG 3640  
DB |||||  
QY 3911 AAAGGGGGAATGAAGACCCCACTGTAGGTTTGGCAGCTAGTCTTAAGTAAAGCCATT 3970  
DB |||||  
QY 3641 ATCCGCGTGAACAAAGGGCGCCCCCTGCGCTTCCGATCTCTGAGCCCGCCCTTC 3700  
DB |||||  
QY 3971 TTGCAAGGCA-----TGAAATAATACATACTGAGATAGAG 4006  
DB |||||  
QY 3701 CAGTACGGCAACCGCACTTCAACAGTACCCCGAGGACATCAGGACTTCTTCTATCCAG 3760  
DB |||||  
QY 4007 AAGTTCAGATCAAGGTTAGGAACAGAGAGACAGCAGAAATATATGGGCCAAACAGGATATCTG 4066  
DB |||||

Db 3761 AGCTTCCCGCGGCTTGTGTACGAGCGCACCCCTGCGCTACGAGCGCGCGCTGGTG 3820  
Qy 4067 TGG-----TAAACAGTTCCTGCCCGCTCAGGCGCAAGACAGTTGGAACA 4112  
Db 3821 GAGATCCGAGCGACATCAACCTGATCGAGGAGATGTTCTGTATCCGCGTGGAGTACAAG 3880  
Qy 4113 GGAGAAATATGGGCCAACAGGATATCTGTGGTAAGCAGTTCCTGCCCGCGCTCAGGGCCA 4172  
Db 3881 GCGCGAACTTCCCAACAGACCGCCCGTGTATGAGAGACCATCACCGGCCCTGACGCC 3940  
Qy 4173 AGAACAGATGGTCCCGATGGGTCCCGCCCTCAGCAGTTTCTAGAGAAC-----4223  
Db 3941 AGCTTCGAGGTGTGTATCATGAACGACGCGGTGCTGTGGGCCAGGTGATCTCGTGTATC 4000  
Qy 4224 -----CATCAGATGTTTCAGGTTTCAGGTTGCCAGGAGCTGAAATGAC 4263  
Db 4001 CGCTGAACAGCGGCAAGTTCTACAGCTGCCATGCGCACCTGATGAAGAGCAAGGCG 4060  
Qy 4264 CCGTGGCTTTATTTGAACCTA--ACCAATCAGTTTCGCTTCTCGCTTCTGTTCCGCGCTTC 4321  
Db 4061 GTGGTAAGGACTTCCCGAGTACCACTTCATCCAGCACCGCTGGAGAAAGACTTACGTG 4120  
Qy 4322 TGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGCGCGC-----GC 4369  
Db 4121 GAGGACGGCGGCTTCGTGGAGCAGCAGAGACCGCCATCGCCAGCTGACCGCTGGGC 4180  
Qy 4370 CAGTCTCCGATAGACTCGGTGCGCCGGGTACCGGTGTTCTCAATAAACCTTCTGCAGT 4429  
Db 4181 AAGCCCTCGGCGAGCTGACGAGTGGGTGTAACTCGAGCGCGCGATCCGAGTTCTTTC 4240  
Qy 4430 TGATCCGACTCGTGTCTCGTGTCTTGGGAGGGTCTCT-----CTGA 4475  
Db 4241 TGAGCGGAGCTCTGGGGTTCGATAAAATAAAGATTTTATTTAGTCTCCAGAAAAGGGG 4300  
Qy 4476 GTGATTGACTACCCGTCAGCGGGTCTTTCAGTTCCTCCACCTCAACAGGTCTCACTAA 4535  
Db 4301 GGNATGAAGAGACCCACCTGTAGTTTGGCAAGCTAGCTTAAGTAACGCCAATTTTCAAG 4360  
Qy 4536 CATTCCTGATGTGC--CGAGGAGCTCCGTGAGCCGGTTTTTTGTTATAATAAATGCA 4593  
Db 4361 GCATGAAAATAATACATAAATGAGAAATAGAGAAATTCAGATCAAGGTGAGGAAACAGATGA 4420  
Qy 4594 AGAACAGTGTCCCTTCAAGCAGACTACATCTGACTCTCGGCTTTATAAAGAATGTT 4653  
Db 4421 ACAGCTGAATATGGGCCAAACAGGATATCTGTGGTAAGCAGTTCTGCCCCGGCTCAGGG 4480  
Qy 4654 GAAGGGCTCTGTGACTATCTGCCACACGACTTTTAAAGATTTTATG-----4701  
Db 4481 CCAAGAACAGATGGAACAGCTGAATATGGCCAAACAGGATATCTGTGGTAAGCAGTTCC 4540  
Qy 4702 CTTCTGGATGAGGATTTAGTCAATCTATCTGCTGCTATTTTGTGGTCTTCTCGGTATT 4761  
Db 4541 TGCCCCGGCTCAGGGCCAAAGAACAGATGTTCCAGATGCGGTCCAGCCCTCAGCAGTTT 4600  
Qy 4762 TTAATTTT-----4770  
Db 4601 CTAGAGAACATCAGATGTTTCCAGGGTGCCTCAAGGACCTGAAATGACCTGTGCTTCA 4660  
Qy 4771 -----TAGTTTGACTCCCTTCTGAGAGCAGCGGATTCGACAGTAGTT 4815  
Db 4661 TTTGAATCAACCAATCAGTTCTGCTTCTGCTTCTGTTCCGCGCTTCTGCTCCCCGAGCT 4720  
Qy 4816 AATACTCTGAGGCGAGGCTTCTGTGAAAAGGTTGCTGGGCTCAGGTGAGATTTTGGCA 4875  
Db 4721 CAATAAAGAGCCCAACACCCCTCACTCGGGGCGCCAGTCTCCGATGACTGAGTCCGCC 4780  
Qy 4876 TAAAAAGGGTCTGCGCCCTGTGTACAGACAGATCGAATCTAGAGTGCATCTCAGACT 4935  
Db 4781 CCGGTACCCGTGTATCCAATAAACCCCTCTTTCAGTTGTCATCCGACTTGTGGTCTCGCTGT 4840  
Qy 4936 CCCCCGGTTCGGGGCTCTGA-----4957

Db 4841 TCCTTTGGAGGGTCTCTCTGAGTGATTTGACTACCCGTGAGCGGGGTCTTTTCATTTGGG 4900  
Qy 4958 -----TCTCAGGCGCATCTTTTGTCTAGAGATCTCTACGCGCGAC--GCATCTGTG 5004  
Db 4901 GGCTCGTCCGGATCCGGAGAGCCCTGCGAGGACACCGACCCACACCGGAGGTAA 4960  
Qy 5005 GCGGGTACCAGAGCTCGAATTCGTAAATCATGGTCAATAG--CTGTTTCTGTGTGAATGTT 5063  
Db 4961 GCTGCTGCTCGCGCGTTCGGTGTATGACGGTGAACCTCTGACACATGACGCTCCCG 5020  
Qy 5064 TATCGCTCAATTTCCACACAAATACGAGCGCGAAGCATAAAGTGTAAAGCTTGGGT 5123  
Db 5021 GAGACGGTCAACAGCTTGTCTGTAAAGCGATGCGCGGAGCAGCAAGCCGTCAGGCGCG 5080  
Qy 5124 GCCTAATCAGT-----GAGCTAACTCA 5145  
Db 5081 TACGCGGGTGTGGCGGGGTGTCGGGGCGCAGCCATGACCCAGTACGATAGCGATAGCGA 5140  
Qy 5146 CATTAATTCGGTTCGGCTCACTGCCCGCTTTTCAGTTCGGGAAACCTGTGTGCCAGCTGC 5205  
Db 5141 GTGTATACCTGGCTTAACTATGCGCATCAGAGCAGATTGTACTGAGAGTGCAACCATATGC 5200  
Qy 5206 ATTAATGAATCGGCCCAACGCGCGGGAGAGCGGTTTGGGTATTGGGCGCTCTTCGGCTT 5265  
Db 5201 GGTGTGAATAACCGCACAGATGCGTAAAGGAGAAAAATACCGCATCAGGCGCTCTTCGGCTT 5260  
Qy 5266 CCGCTCAGTCACTGACTCGCTGCGTTCGGTTCGGTTCGGCGAGCGGTATCAGTCACT 5325  
Db 5261 CCGCTCAGTCACTGACTCGCTGCGTTCGGTTCGGTTCGGCGAGCGGTATCAGTCACT 5320  
Qy 5326 CAAAGCGCGTAAATACGGTTATCCACAGAAATCAGGGGATAAAGCGCAGGAAAGAACATGTGAG 5385  
Db 5321 CAAAGCGCGTAAATACGGTTATCCACAGAAATCAGGGGATAAAGCGCAGGAAAGAACATGTGAG 5380  
Qy 5386 CAAAGCGCGCAGCAAAAGCGCAGAAACCGTAAAGCGCGGTGCTGCGGTTTTTCCATA 5445  
Db 5381 CAAAGCGCGCAGCAAAAGCGCAGAAACCGTAAAGCGCGGTGCTGCGGTTTTTCCATA 5440  
Qy 5446 GCGTCCGCGCCCTGACGAGCATCAAAATCAGAGCTCAAGTCAGAGGTGCGGAAACC 5505  
Db 5441 GCGTCCGCGCCCTGACGAGCATCAAAATCAGAGCTCAAGTCAGAGGTGCGGAAACC 5500  
Qy 5506 CGACAGGACTATAAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTCGT 5565  
Db 5501 CGACAGGACTATAAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTCGT 5560  
Qy 5566 TTCGACCTGCGCTTACCGGATACCTGTGCGCTTTCCTCCGCGGAGCGTGGCGC 5625  
Db 5561 TTCGACCTGCGCTTACCGGATACCTGTGCGCTTTCCTCCCTTCGGGAAAGCGTGGCGC 5620  
Qy 5626 TTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTTCAGTTCGGTTCAGTTCGGTTC 5685  
Db 5621 TTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTTCAGTTCGGTTCAGTTCGGTTC 5680  
Qy 5686 GCTGTGTGACGAAACCCCGCTTACGCGCGACCGCTTTCGGCTTTCAGTTCGGTTCAGTTC 5745  
Db 5681 GCTGTGTGACGAAACCCCGCTTACGCGCGACCGCTTTCGGCTTTCAGTTCGGTTCAGTTC 5740  
Qy 5746 TTGAGTCCAAACCGGTAAAGACACGACTTATCGCCACTGGCAGCAGCACTGTGTAAACAG 5805  
Db 5741 TTGAGTCCAAACCCCGGTAAAGACACGACTTATCGCCACTGGCAGCAGCACTGTGTAAACAG 5800  
Qy 5806 TTGAGCAGCAGGATGTTAGGCGGTGCTACAGGTTCTTGAAGTGGTGGCTTAACAG 5865  
Db 5801 TTGAGCAGCAGGATGTTAGGCGGTGCTACAGGTTCTTGAAGTGGTGGCTTAACAG 5860  
Qy 5866 GCTACACTAGAAAGGACAGTATTGTTGTTATCTGCGCTCTGTGAAGCGCAGTTACTCTCGAA 5925  
Db 5861 GCTACACTAGAAAGGACAGTATTGTTGTTATCTGCGCTCTGTGAAGCGCAGTTACTCTCGAA 5920  
Qy 5926 AAAAGAGTTGGTGTAGCTTTGATCCGGCAACAAACCAACCGCTGTGAGCGGTGTTTTTTG 5985  
Db 5921 AAAAGAGTTGGTGTAGCTTTGATCCGGCAACAAACCAACCGCTGTGAGCGGTGTTTTTTG 5980



QY 5986 TTTCACGACGAGATTAGCGCAGAAAAAAGGATCTCAAGAGATCCTTTGATCTTTT 6045  
DB 5981 TTTCACGACGAGATTAGCGCAGAAAAAAGGATCTCAAGAGATCCTTTGATCTTTT 6040  
QY 6046 CTACGGGCTCTGACGCTCAGTGGAAACGAAACTCAGCTTAAGGGATTTTGGTCATGAGAT 6105  
DB 6041 CTACGGGCTCTGACGCTCAGTGGAAACGAAACTCAGCTTAAGGGATTTTGGTCATGAGAT 6100  
QY 6106 TATCAAAAGGATCTTACCTAGATCTCTTTTAAATTTAAAAATGAAGTTTAAATCAATCT 6165  
DB 6101 TATCAAAAGGATCTTACCTAGATCTCTTTTAAATTTAAAAATGAAGTTTAAATCAATCT 6160  
QY 6166 AAAGTATATATGAGTAAACTTGGTCTGACAGTGTACCAATGCTTAATAGTGAGGACCTA 6225  
DB 6161 AAAGTATATATGAGTAAACTTGGTCTGACAGTGTACCAATGCTTAATAGTGAGGACCTA 6220  
QY 6226 TCTCAGCGATCTGTCTATTTCGTTCCATCCATAGTTTGCCTGACTCCCGCTCGTGTAGATAA 6285  
DB 6221 TCTCAGCGATCTGTCTATTTCGTTCCATCCATAGTTTGCCTGACTCCCGCTCGTGTAGATAA 6280  
QY 6286 CTACGATACGGAGGCTTACCATCTGGCCCGCAGTCTGCAATGATACCGGAGACCCAC 6345  
DB 6281 CTACGATACGGAGGCTTACCATCTGGCCCGCAGTCTGCAATGATACCGGAGACCCAC 6340  
QY 6346 GCTCACCGGCTCCAGATTATCAGCAATAAACCCAGCCAGCCGGAAGGCGCGAGCGAGAA 6405  
DB 6341 GCTCACCGGCTCCAGATTATCAGCAATAAACCCAGCCAGCCGGAAGGCGCGAGCGAGAA 6400  
QY 6406 GTGTGCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAAATTTGTCGGGAGAGCTAGAG 6465  
DB 6401 GTGTGCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAAATTTGTCGGGAGAGCTAGAG 6460  
QY 6466 TAAAGTAGTTCCGACGTTAATAGTTTGGCGCAAGTTGTCGCAATGCTACAGC-TCGTGG 6524  
DB 6461 TAAAGTAGTTCCGACGTTAATAGTTTGGCGCAAGTTGTCGCAATGCTACAGC-TCGTGG 6520  
QY 6525 TGTCACGCTCGTCTGTTGGTATGCTTCATTACAGCTCCGGTTCCCAACGATCAAGCGAG 6584  
DB 6521 TGTCACGCTCGTCTGTTGGTATGCTTCATTACAGCTCCGGTTCCCAACGATCAAGCGAG 6580  
QY 6585 TTACATGATCCCGATGTTGCAAAAAAGGTTAGTCTCTCGGTCCTCGATCGTTG 6644  
DB 6581 TTACATGATCCCGATGTTGTTGCAAAAAAGGTTAGTCTCTCGGTCCTCGATCGTTG 6640  
QY 6645 TCAGAAGTAAGTTGCGCGAGTGTATCACTCATGTTATGCGACACTGCGATATCTC 6704  
DB 6641 TCAGAAGTAAGTTGCGCGAGTGTATCACTCATGTTATGCGACACTGCGATATCTC 6700  
QY 6705 TTACTGTCTATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCAT 6764  
DB 6701 TTACTGTCTATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCAT 6760  
QY 6765 TCTGAGATAGTGTATGCGGACCGAGTGTCTTGTGCGCGGCTCAATACGGGATATA 6824  
DB 6761 TCTGAGATAGTGTATGCGGACCGAGTGTCTTGTGCGCGGCTCAATACGGGATATA 6820  
QY 6825 CCGCGCCACATAGCAGACTTTAAAGTGTCTATCATTTGAAAGGTTCTTCGGGCGAA 6884  
DB 6821 CCGCGCCACATAGCAGACTTTAAAGTGTCTATCATTTGAAAGGTTCTTCGGGCGAA 6880  
QY 6885 AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCAGTTGAAACCCACTCGTGACCCCA 6944  
DB 6881 AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCAGTTGAAACCCACTCGTGACCCCA 6940  
QY 6945 ACTGATCTTCAGCATCTTTTACTTTTCAACAGGTTTCTGGGTGAGCAAAAAACAGGAGGC 7004  
DB 6941 ACTGATCTTCAGCATCTTTTACTTTTCAACAGGTTTCTGGGTGAGCAAAAAACAGGAGGC 7000  
QY 7005 AAAATGCCGCAAAAAAGGAAATAGCGGACACGGAATGTTGAATACTACTACTCTTCC 7064  
DB 7001 AAAATGCCGCAAAAAAGGAAATAGCGGACACGGAATGTTGAATACTACTACTCTTCC 7060

QY 7065 TTTTTCATATTATTGAAGCAATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTG 7124  
DB 7061 TTTTTCATATTATTGAAGCAATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTG 7120  
QY 7125 AATGTATTATTGAAAAATAAACAAATAGGGTTTCGCGCACATTTTCCCGAAAAAGTGCCAC 7184  
DB 7121 AATGTATTATTGAAAAATAAACAAATAGGGTTTCGCGCACATTTTCCCGAAAAAGTGCCAC 7180  
QY 7185 CTGACGCTTAAGAAACCAATTTATTATCATGATTAACCTATAAAAAATAGGCGTATCAGA 7244  
DB 7181 CTGACGCTTAAGAAACCAATTTATTATCATGATTAACCTATAAAAAATAGGCGTATCAGA 7240  
QY 7245 GGCCCTTTCGTCT 7257  
DB 7241 GGCCCTTTCGTCT 7253

RESULT 11  
AAT13393  
ID AAT13393 standard; DNA; 5364 BP.  
AC AAT13393;  
XX AC  
XX 25-MAR-2003 (revised)  
DT 24-JUN-1996 (first entry)  
XX  
XX Hybrid vector pM1.  
XX  
KW Hybrid; vector; gene transfer; gene therapy; haematopoietic stem cell;  
KW retroviral; murine embryonic stem cell virus; MESV;  
KW Moloney murine sarcoma virus; (MoMuSV); Friend murine leukaemia virus;  
KW F-MuLV; ds.  
XX  
OS Synthetic.  
XX  
FN DE19503952-A1.  
XX  
PD 14-MAR-1996.  
XX  
PF 07-FEB-1995; 95DE-01003952.  
XX  
PR 08-SEP-1994; 94DE-04431973.  
XX  
XX (BOEF) BOEHRINGER MANNHEIM GMBH.  
XX  
PI Baum C, Stocking-Harbers C, Ostertag W;  
XX  
DR WPI; 1996-152306/16.  
XX  
PT Hybrid retroviral vectors - for gene transfer into haematopoietic stem  
PT cells.  
PS  
PS Claim 10; Page 21-25; 42pp; German.  
XX  
XX New hybrid vectors comprise (1) a leader region including the U5 region  
CC and tRNA primer binding site of murine embryonic stem cell virus (MESV)  
CC or Moloney murine sarcoma virus (MoMuSV), and (2) a 3'-LTR including the  
CC U3 and R regions of a Friend murine leukaemia virus (F-MuLV). The vectors  
CC are useful for ex-vivo or in-vivo gene therapy. High levels of gene  
CC transfer can be achieved in haematopoietic stem cells and their myeloid  
CC (non-lymphatic) progeny. pSF1, pSF2, pSF3 and pM1 (sequences given in  
CC AAT13390-T13393) are examples of such vectors. Vector pSF-MDR (sequence  
CC given in AAT13394) is based on the MESV vector R224. (Updated on 25-MAR-  
CC 2003 to correct PR field.)  
XX  
SQ Sequence 5364 BP; 1233 A; 1454 C; 1355 G; 1322 T; 0 U; 0 Other;  
Query Match 40.6%; Score 3111; DB 2; Length 5364;  
Best Local Similarity 92.0%; Pred. No. 0;  
Matches 3414; Conservative 0; Mismatches 210; Indels 87; Gaps . 9;  
QY 3868 ACGAGCCATAGATAAAATAAGATTTTATTAGTCTCCAGAAAAAGGGGGAATGAAG 3927  
|||||

Db 1655 ACAGAGCATAGATAGAAATTTATTTAGTCTCCAGAAAAAGGGGGGAATGAAG 1714  
 QY 3928 ACCCCACCTGTAGTTTGGCAAGC-----TAGCTTAAGTAAGCCCAT 3969  
 Db 1715 ACCCCACCTGTAGTTTGGCAAGCTAGAGTCGCTTAGCCTGATAGCGGAGTAAGCCCAT 1774  
 QY 3970 TTTGCAAGCATGG-AAAAATACATAAATGAGATAGAGAAAGTTTCAGATCAAGGTTTAGGAA 4028  
 Db 1775 TTTGCAAGCATGGAAAAATACCAACCAAGAAATAGGAAGTTTCAGATCAAGGGGGTA 1834  
 QY 4029 CAGAGAGACAGCAGAAATATGGGCAAAACAGGATATCTGTGTAAGCAGTTCTTCGCCCGC 4088  
 Db 1835 CATGAAATAGCTTAAAGTGGGCAAAACAGGATATCTGGGTGAGCAGTTTCGGCCCCGG 1894  
 QY 4089 TCAGGSCCAAGAACAGTCGGAACAGGAGATATGGGCCAAAC-----AGGATATCTG 4140  
 Db 1895 CCGGGGCAAGAACAGATGGTACCGCAGTTTCGGCCCCGGCCCGGGGCCAAGAACAGATG 1954  
 QY 4141 TGGTAAGCAGTTTCTCGCCCGGCTCAGGGCCCAAGAACAGATGGTCCCCAGATGCGGTCCC 4200  
 Db 1955 GTCACCGCAGTTTCGGCCCCGGCCCGGGGCCAAGAACAGATGGTCCCCAGATGAGCCCA 2014  
 QY 4201 GGCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCCCCAAGGACCTGAAT 4260  
 Db 2015 TCCCTCAGCAGTTTCTTAAAGACCCATCAGATGTTTCCAGGCTCCCCCAAGGACCTGAAT 2074  
 QY 4261 GACCTGTGCTTATTTGAATTAACCAATCAGTTTCGCTTCTGCTTCTGTCGCGCTT 4320  
 Db 2075 GACCTGTGCTTATTTGAATTAACCAATCAGCTTCTGCTTCTGCTTCTGTCGCGCTT 2134  
 QY 4321 CTGCTCCCGAGCTCAATAAAAGAGCCCAACCCCTCACTCGGGCGCCAGTCTCTCCGA 4380  
 Db 2135 CTGCTCCCGAGCTCTAATAAGAGCTCAACCCCTCACTCGGGCGCCAGTCTCTCCGA 2194  
 QY 4381 TAGACTGCGTGGCCGGGTACCGGTCTCTCAATAAACCTCTTGAGTTGATCGGACT 4440  
 Db 2195 TTGACTGAGTCGCGCGGTACCGGTCTCTCAATAAACCTCTTGAGTTGATCGGACT 2254  
 QY 4441 CGTGTCTGCTTCTTGGAGGGGTCT-CTCTGAGTCAATGACTACCCGTCAGC-GGG 4498  
 Db 2255 CGTGTCTGCTTCTTGGAGGGGTCTCTCTGAGTCAATGACTACCCGTCAGCGGG 2314  
 QY 4499 GTCTTTTCACTTCTCCCACTACACAGGTCTCACTAAACATCTCTGATGTCGCGCAGGGAC 4558  
 Db 2315 GTCTTTTCACTTCTCCCACTACACAGGTCTCACTAAACATCTCTGATGTCGCGCAGGGAC 2374  
 QY 4559 TCGGTCAAGCCGGTTTTGTTATATAAATAAATGCAAGACAGTTTCCCTTCAAGCCAGA 4618  
 Db 2375 TCGGTCAAGCCGGTTTTGTTATATAAATAAATGCAAGACAGTTTCCCTTCAAGCCAGA 2434  
 QY 4619 CTACATCTGACTCTCGGCTTTATAAAGAAATGTTGAAGGGCTCTGTGAGTATCTGCCA 4678  
 Db 2435 CTACATCTGACTCTCGGCTTTATAAAGAAATGTTGAAGGGCTCTGTGAGTATCTGCCA 2494  
 QY 4679 CACGACTTTTAAAGATTTTATGCTTCTGAGTGAAGGATTTAGTCAATCTCTCGTC 4738  
 Db 2495 CACGAC-TTTTAAAGATTTTATGCTTCTGAGTGAAGGATTTAGTCAATCTCTCGTC 2553  
 QY 4739 TATTTGCTGCTTCTCGTATTTTAAATTTCTAGTTTGCATCTCCCTTCTCAGAGCAG 4798  
 Db 2554 TATTTGCTGCTTCTCGTATTTTAAATTTCTAGTTTGCATCTCCCTTCTCAGAGCAG 2613  
 QY 4799 GCGATTGCAAGTAGTTAATACTCTGAGGGCAGGCTTCTGTGAAGAGTTGCTGGGCTC 4858  
 Db 2614 GCGATTGCAAGTAGTTAATACTCTGAGGGCAGGCTTCTGTGAAGAGTTGCTGGGCTC 2673  
 QY 4859 AGTGTGAGATTTTGCATATAAAGGGGTCTCGCCCTGTGTACAGACAGATCGGAATCTA 4918  
 Db 2674 AGTGTGAGATTTTGCATATAAAGGGGTCTCGCCCTGTGTACAGACAGATCGGAATCTA 2733  
 QY 4919 GAGTGATATCTCAGAGTCCCGGTTCCGGGCTCTGATCTCAGGGCATCTTTGCTAG 4978  
 Db 2734 GAGTGATATCTCAGAGTCCCGGTTCCGGGCTCTGATCTCAGGGCATCTTTGCTAG 2793

QY 4979 AGATCTCTTACGCCGAGCATCTGTGGCGGGTACCGAGTCTGAATTTGTTATCATGTGTC 5038  
 Db 2794 AGATCTCTTACGCCGAGCATCTGTGGCGGGTACCGAGTCTGAATTTGTTATCATGTGTC 2853  
 QY 5039 ATAGCTGTTTC-----CTGTGTGAAATTTGTTATTCGGTCTCAAAATTCACACA 5085  
 Db 2854 GGGCTTATATCGCCGACATCACCGATGGGAAGATCGGGCTCGCACATTCGGGCTCATGA 2913  
 QY 5086 ACATAGAGCGGGAAGCATTAAGTGTAAAGCTTGGGGTGCCTTAATGAGT----- 5134  
 Db 2914 GCGCTTGTTCGGCTGGGTATGTTGGCAGGCCCTCGCGGGGAGCTGTTGGGCGCA 2973  
 QY 5135 -----GAGTAACTCACATTAATTTGCGTTGCGC 5162  
 Db 2974 TCTCTTGATGACCATTTCTTGGCGGGCGGTCTCAAGGCTCTCACTACTACTG 3033  
 QY 5163 TCATGCGCGCTTTCAGTTCGGGAAACCTGTGTCAGCTGATTAATGAATCGGCCAA 5222  
 Db 3034 GCTGCTTCTTAATGACGAGTTCGCATAAGGAGAGCGTCTGCAITTAATGAATCGGCCAA 3093  
 QY 5223 CCGCGGGGAGAGCGGTTTGGGTATTTGGGCGCTCTTTCGGCTTCTCGTCACTGACTCG 5282  
 Db 3094 CCGCGGGGAGAGCGGTTTGGGTATTTGGGCGCTCTTTCGGCTTCTCGTCACTGACTCG 3153  
 QY 5283 CTGCGCTCGGCTGCTGCGGAGCGGTATCAGCTCACTCAAAAGCGGTAAATACGG 5342  
 Db 3154 CTGCGCTCGGCTGCTGCGGAGCGGTATCAGCTCACTCAAAAGCGGTAAATACGG 3213  
 QY 5343 TTATCCACAGATCAGGGGATAACCGCAGGAAGAAATCATGTGAGCAAAAGGCCAGAAAAG 5402  
 Db 3214 TTATCCACAGATCAGGGGATAACCGCAGGAAGAAATCATGTGAGCAAAAGGCCAGAAAAG 3273  
 QY 5403 GCAGAAACCGTAAAGGCGGTTGCTGGGTTTTTCCATAGCTCGGCGCCCTGAC 5462  
 Db 3274 GCAGAAACCGTAAAGGCGGTTGCTGGGTTTTTCCATAGCTCGGCGCCCTGAC 3333  
 QY 5463 GAGCATCAAAATTCGACGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAAGA 5522  
 Db 3334 GAGCATCAAAATTCGACGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAAGA 3393  
 QY 5523 TACAGGGGTTTTCCCTCGGAAGTCCCTCGTGGCTCTCTCTGTTCCGACCTGCGGCTT 5582  
 Db 3394 TACAGGGGTTTTCCCTCGGAAGTCCCTCGTGGCTCTCTCTGTTCCGACCTGCGGCTT 3453  
 QY 5583 ACCGATACGTGTCGCTTCTCCCTTCGGGAAGCGTGGGCTTCTCATAGCTCAGCG 5642  
 Db 3454 ACCGATACGTGTCGCTTCTCCCTTCGGGAAGCGTGGGCTTCTCAATGCTCAGCG 3513  
 QY 5643 TGTAGTATCTCAGTTCCGTTGAGTCTGCTCCAAAGCTGGGTGTGTGACGAAACC 5702  
 Db 3514 TGTAGTATCTCAGTTCCGTTGAGTCTGCTCCAAAGCTGGGTGTGTGACGAAACC 3573  
 QY 5703 CCGGTTACGCGGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAAGTCAACCGGTA 5762  
 Db 3574 CCGGTTACGCGGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAAGTCAACCGGTA 3633  
 QY 5763 AGACAGATTTATCGGCTGCGGACGACCTGTGTAAAGGATTTAGCAGCGAGGAT 5822  
 Db 3634 AGACAGATTTATCGGCTGCGGACGACCTGTGTAAAGGATTTAGCAGCGAGGAT 3693  
 QY 5823 GTAGGCGGTGTACAGAGTTCTTGAAGTGGTGGCTTAACGCTCACTAGAAAGGACA 5882  
 Db 3694 GTAGGCGGTGTACAGAGTTCTTGAAGTGGTGGCTTAACGCTCACTAGAAAGGACA 3753  
 QY 5883 GTATTTGTTATCTGCGCTCTGCTGAAGCAGTTTACTCTCGGAAAGAGTTTGGTAGCTCT 5942  
 Db 3754 GTATTTGTTATCTGCGCTCTGCTGAAGCAGTTTACTCTCGGAAAGAGTTTGGTAGCTCT 3813  
 QY 5943 TGATCCGGCAAAACCAACCGCTGGTAGCGGTGTTTTTTTTTTGTTTGAAGCAGAGATT 6002  
 Db 3814 TGATCCGGCAAAACCAACCGCTGGTAGCGGTGTTTTTTTTTTGTTTGAAGCAGAGATT 3873

Qy	6003	ACGGCGCAGAAAAAGGATCTCAAGAGAGATCCTTTTGATCTTTTCTA CGGGGTCTGACGCT	6063
Db	3874	ACGGCGCAGAAAAAGGATCTCAAGAGAGATCCTTTTGATCTTTTCTA CGGGGTCTGACGCT	3933
Qy	6063	CAGTGGAAACGAAAACTCACGTTAAAGGGATTTTGTCTCATGAGATTATCAAAAAAGATCTTC	6122
Db	3934	CAGTGGAAACGAAAACTCACGTTAAAGGGATTTTGTCTCATGAGATTATCAAAAAAGATCTTC	3993
Qy	6123	ACCTAGATCTCTTTTAAATTTAAAAATGAAGTTTAAAACTAAATCTAAAGTATATATGAGTAA	6182
Db	3994	ACCTAGATCTCTTTTAAATTTAAAAATGAAGTTTAAAACTAAATCTAAAGTATATATGAGTAA	4053
Qy	6183	ACTTTGGTCTGACAGCTTACCAATGCTTAAATCAGTGAGGCACTTATCTCAGCGATCTGTCCTA	6242
Db	4054	ACTTTGGTCTGACAGCTTACCAATGCTTAAATCAGTGAGGCACTTATCTCAGCGATCTGTCCTA	4113
Qy	6243	TTTTCGTTTCATCCATAGTTTGCCTGACTCCCGTCCGTGTAGATAACTACGATACGGGAGGGC	6302
Db	4114	TTTTCGTTTCATCCATAGTTTGCCTGACTCCCGTCCGTGTAGATAACTACGATACGGGAGGGC	4173
Qy	6303	TTACCATCTTGGCCCCAGTGTGCAATGATACCGCGAGACCCACGCTCACCGGTCTCCAGAT	6362
Db	4174	TTACCATCTTGGCCCCAGTGTGCAATGATACCGCGAGACCCACGCTCACCGGTCTCCAGAT	4233
Qy	6363	TTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAGTGTCTGCAACTTTA	6422
Db	4234	TTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAGTGTCTGCAACTTTA	4293
Qy	6423	TCCGCCTCCATCCAGTCTATTATATGTTGTCGGGAAGCTAGATTAAGTATTGCCCAGTT	6482
Db	4294	TCCGCCTCCATCCAGTCTATTATATGTTGTCGGGAAGCTAGATTAAGTATTGCCCAGTT	4353
Qy	6483	AAATAGTTTGGCCAACTGTTTGGCCATTTGCTACAGGC--TCGTGGTGTCAACGCTCGTCTGTT	6541
Db	4354	AAATAGTTTGGCCAACTGTTTGGCCATTTGCTACAGGCATCGTGGTGTCAACGCTCGTCTGTT	4413
Qy	6542	GGTATGCTCTCATTCAGCTCCGPTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATG	6601
Db	4414	GGTATGCTCTCATTCAGCTCCGPTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATG	4473
Qy	6602	TTGTGCAAAAAAGCGGTAGCTCTTCGGTCTCCCGATCGTGTGAGAACTAAGTTGGCC	6661
Db	4474	TTGTGCAAAAAAGCGGTAGCTCTTCGGTCTCCCGATCGTGTGAGAACTAAGTTGGCC	4533
Qy	6662	GCAGTGTATCACTCATGCTGTTATGGCAGCACTGCATAATTTCTCTTACTGTCAATGCCATCC	6721
Db	4534	GCAGTGTATCACTCATGCTGTTATGGCAGCACTGCATAATTTCTCTTACTGTCAATGCCATCC	4593
Qy	6722	GTAAAGTCTTTTCTGACTGCTGAGTACTCAACCAAGTCATTCTGAGAAATAGTGTATG	6781
Db	4594	GTAAAGTCTTTTCTGACTGCTGAGTACTCAACCAAGTCATTCTGAGAAATAGTGTATG	4653
Qy	6782	CGGGCAACGAGTTGCTCTTCCCGCGCTCAATACGGGATTAATACGGGCCACATAGCAGA	6841
Db	4654	CGGGCAACGAGTTGCTCTTCCCGCGCTCAATACGGGATTAATACGGGCCACATAGCAGA	4713
Qy	6842	ACTTTAAAAAGTCTCATCATTTGAAAAACGTTCTTTCGGGGCGAAAACTCTCAAGGATCTTA	6901
Db	4714	ACTTTAAAAAGTCTCATCATTTGAAAAACGTTCTTTCGGGGCGAAAACTCTCAAGGATCTTA	4773
Qy	6902	CCGCTGTTGAGATCCAGTTTCGATGTAAACCACTCGTGACCCCACTGATCTTTCAGCATCT	6961
Db	4774	CCGCTGTTGAGATCCAGTTTCGATGTAAACCACTCGTGACCCCACTGATCTTTCAGCATCT	4833
Qy	6962	TTTACTTTTCCACGCGTTTCTGGGTGAGCAAAAAACAGGAAAGGCAAAATGCGCAAAAAAG	7021
Db	4834	TTTACTTTTCCACGCGTTTCTGGGTGAGCAAAAAACAGGAAAGGCAAAATGCGCAAAAAAG	4893
Qy	7022	GGAAATAGGCGCACACGGAAATGTTGAAATACTCATACTCTTCTCTTTTCAATATTATTTGA	7081
Db	4894	GGAAATAGGCGCGACACGGAAATGTTGAAATACTCATACTCTTCTCTTTTCAATATTATTTGA	4953
Qy	7082	AGCATTTTATCAGGGTTATTGCTCTCATGAGCGGATACATATTTGAAATGATTTTAGAAAAAT	7141

Db	4954	AGCATTTTATCAGGGTTATTGCTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAAT	5013
Qy	7142	AAACAAATATAGGGTTCCGCGCACATTTCCCGGAAAGTGCACCTGACGTCTAAGAAACC	7201
Db	5014	AAACAAATATAGGGTTCCGCGCACATTTCCCGGAAAGTGCACCTGACGTCTAAGAAACC	5073
Qy	7202	ATTATTATCATGACATTAACTTATAAAATAGGCGTATCACGAGGCCCTTTTCGTCTCGCG	7261
Db	5074	ATTATTATCATGACATTAACTTATAAAATAGGCGTATCACGAGGCCCTTTTCGTCTCGCG	5133
Qy	7262	CGTTTCGGTGATGACGGTGAAAACTCTGACATGACGTCCCGAGACGGTCAAGCT	7321
Db	5134	CGTTTCGGTGATGACGGTGAAAACTCTGACATGACGTCCCGAGACGGTCAAGCT	5193
Qy	7322	TGCTCTGAAGCGGATGCGGGAGCAGACAGCCCGTCAAGGCGCGTCAAGCGGCTGTTGGC	7381
Db	5194	TGCTCTGAAGCGGATGCGGGAGCAGACAGCCCGTCAAGGCGCGTCAAGCGGCTGTTGGC	5253
Qy	7382	GGGTGTCGGGGCTGGCTTAACTATGGCGCATCAGAGCAGATTGTACTGAGAGTGCACCAT	7441
Db	5254	GGGTGTCGGGGCTGGCTTAACTATGGCGCATCAGAGCAGATTGTACTGAGAGTGCACCAT	5313
Qy	7442	ATGCGGTGTGAATATACCGCACAGATCGGTAAAGAGAAAAATACGCGATCAGG	7492
Db	5314	ATGCGGTGTGAATATACCGCACAGATCGGTAAAGAGAAAAATACGCGATCAGG	5364
RESULT 12			
AA167595			
ID AA167595 standard; DNA; 6444 BP.			
XX			
AC	AA167595;		
XX			
DT	27-FEB-2002 (first entry)		
XX			
DE	Nucleotide sequence of a cloning vector pLXRN.		
XX			
KW	Hexokinase; cell proliferation; glycolytic tumour; cancer; mutant;		
KW	type II hexokinase; cytostatic; glycolysis inhibitor; gene therapy; ds.		
XX			
OS	Synthetic.		
XX			
FN	W0200168667-A1.		
XX			
PD	20-SEP-2001.		
XX			
PF	14-MAR-2001; 2001WO-US008335.		
XX			
PR	14-MAR-2000; 2000US-0189222P.		
XX			
PA	(UWJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.		
XX			
PI	Pedersen PL, Mathupala SP;		
XX			
DR	WPI; 2002-049005/06.		
XX			
PT	Inhibiting proliferation of highly glycolytic tumors, e.g. in gastric		
PT	cancer, hepatoma, colorectal cancer or lung cancer, by contacting cells		
PT	with antisense molecules that hybridize with a nucleic acid encoding		
XX	hexokinase.		
XX			
PS	Claim 21; Fig 7A-C; 59pp; English.		
XX			
CC	The invention provides a method for inhibiting proliferation of tumour		
CC	cells characterized by having a highly glycolytic phenotype. The method		
CC	involves contacting the cells with an antisense polynucleotide or		
CC	oligonucleotide that hybridizes with a mRNA encoding a hexokinase under		
CC	conditions that allow hybridization of the antisense polynucleotide with		
CC	the mRNA, thus inhibiting the proliferation of tumour cells. The method		
CC	is useful for inhibiting proliferation of highly glycolytic tumours or		
CC	for modulating the expression of a hexokinase in highly glycolytic		
CC	tumours. In particular, the cellular proliferative disorder comprises low		

CC grade astrocytoma, anaplastic astrocytoma, glioblastoma, medulloblastoma,  
CC gastric cancer, hepatoma, colorectal cancer, colorectal adenoma, acute  
CC myelogenous leukemia, lung cancer, renal cancer, leukemia, breast cancer,  
CC prostate cancer, endometrial cancer, bone cancer, squamous cell cancer,  
CC and neuroblastoma. The present sequence represents the nucleotide  
C sequence of a cloning vector pLXRN

Sequence 6444 BP; 1476 A; 1758 C; 1668 G; 1542 T; 0 U; 0 Other;

Query Match 40.5%; Score 3103.4; DB 6; Length 6444;

Best Local Similarity 71.0%; Pred. No. 0;

Matches	4779;	Conservative	0;	Mismatches	1466;	Indels	486;	Gaps	31;
<p> </p>									

QY 549 TTTTGAAGACCCACCTGTAGGTTTGGCAAGCTAGCTTAAGTAACGCCATTTTGCAAGG 608

Db 174 TTTTGAAGACCCACCCGTAGG--TGGCAAGCTAGCTTAAGTAAAGCCCACTTTTGCAAGG 231

609 CATGC-AAAAATACATTAACCTGCAATACACACCTTACCATACCATACAC  
Ov

[illegible]

\_\_\_\_\_

[illegible]

QY / Z / AAGAACAGTGGAAACAGGAGAAI - ATGGGCCCAACACAGGAIAATCTGTGGTAAAGCAGTTCCT / 85

Db 348 AAGAACAGATGAGACAGCTGAGTGATGGGCCAAACAGGATATCTGTGGTAAGCAGTTCCT 407

Qy 786 GCCCGGCTCAGGGCCAAGAACAGATGGTCCCAAGATGGGTCCCGCCCTCAGCAGTTTC 845

**D<sub>b</sub>** 408 GCCCGGCTCGGGGCCAAGAACAGATGGTCCCCAGATGGGTCCAGCCCTCAGCAGTTTC 467

QY 846 TAGAGAACCATCAGATGTTTCCAGGGTGCCCAAGGACCTG-AAATGACCCCTGTGCCTTA 904

468 TAGTGAAATCATCAGATGTTTCCAGGGTGCCCCCAAGGACCTGAAAAATGACCCCTGTACCTTA 527

**Ov**

**905 TTTGAACTAACCAATCAGTTCGGTCTCCCGGCCTTCGTCTGCCAAGCT 964**

20  
 19  
 18  
 17  
 16  
 15  
 14  
 13  
 12  
 11  
 10  
 9  
 8  
 7  
 6  
 5  
 4  
 3  
 2  
 1  
 0

**U**

\_\_\_\_\_

[illegible]

QY 1V23 C6691ACCCG1AT1CCCCATATTAAGGCCCTCTTGC1GTG11TGCA1CCGATACGTGGAC1CG1G 1V84

DB 648 C G G G T A C C C G I A T T C C C A A T A A A G C C T C T T G C T G T T T G C A T C C G A A T C G T G G T C T C G C T G 707

Qy 1085 ATCCTTGGGAGGGTCTCCTCAGATTGATTGACTGCCACCTCGGGGTCTTTCATTGGA 1144

Db 708 TTCCTGGAGGGTCTCCTCTGAGTGATTGACTACCCACGACGGGGTCTTTTCATTTGGG 767

Qy 1145 GGTTCACCGAGATTTTGGAGACCCCTGCCCCAGGGACCAACCGACCCCGCCGGGAGGTA 1204

D<sub>b</sub> 768 GGCTCGTCCGGGATTTGGAGACCCCTGCCAGGGACCACCGA-CCCACCACGGGAGGTA 826

1205 AGCTGGCCAGCAA CTTATCTGTGTCTGTCCGATTGTCTCTA GTGTCTATGA CTGATTTTATG 1264

[illegible]

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

[illegible]

QY 1525 UACGAGTTCGGAACACCCCGGCCCGCAACCCCTGGGAAACGATCCAGGGACATTCGGGGGCCGT 1585

Db 2122 CAATGTAGGGCCCGGAAACCTGGCCCTGTCTTCTTGAGGAGCATTCCTAGGGGTCTTTC 2181  
Qy 2580 TGTAGCTTGTGACACGAGAGAGCAGGTAAGGCTCCAAAGCTGTGATCTACTTGACATC 2639  
Db 2182 -----CCCTCTCGCCCAAGGAATGCAAGGTCTGTGAATGTCTGTAAGGAAGCAGTTC 2234  
Qy 2640 CACCCGGCAGACTGTGTGCGCAAGCAGATTGACGGGTAGCGGTAGCGGTAACCGACTTCAC 2699  
Db 2235 CTCTGGAAGCTTCTTTGAAGACAAACAAAGCTGTGTAGCGACCCCTTTGAGGCGAGCGGAACC 2294  
Qy 2700 CTTCACCATCAGCAGCCTCCAGCCAGAGAGACATCCGCCACTACTACTGCGCAGCAATATAG 2759  
Db 2295 CCCCACTGCGCAGAGGTGCCTCTCGCGCCAAAGCCAGT-----GTATAGATACA 2347  
Qy 2760 CCTCTATCGGTGCTTGGCCAAAGGACCAAGGTGGAATCAAAAGAGGTGCTCAGGATC 2819  
Db 2348 CCTGCAAAAGGGGCACAAACCCAGGTGCCAGGTGTGAGTTGGATAGTTGTGGAAGAGTTC 2407  
Qy 2820 GGGTGGATCCGGCTCTGGTGGCTCAGGATCGGAGTCCAACTGTGTGGAGAGCGGTGGAGG 2879  
Db 2408 AATGGCTCTCTCAAGCGTATGTAACAAAGGGGCTGAAGGATGCCAGAGGTATACCCCAT 2467  
Qy 2880 TGTGTGCAACTGGCCGCTCCCTCGCTGTCTGCTCCGATCTGCTGCTTCCGATTTCCAC 2939  
Db 2468 TGTATGGGATCTGATCTGGGGCTCGGTGCACATGCTTTACATGT-----TTTAG 2519  
Qy 2940 CACATATTTGGATAGTTGGGTGAGACAGGCACCTGGAAGAGTCTTTAGTGGATTTGGAGA 2999  
Db 2520 TCGAGTTAAAGAAAGCTCTAGGCCCCCGAACACACAGGGGACGTGTTTCTTT-----GA 2575  
Qy 3000 AATTCATCCAGATAGCAGTACGATTAATCTATGCGCCGTCTCTAAAGGATAGATTTACAAT 3059  
Db 2576 AAAACACGATATAGCTTTGCCAACAAACCCCGGATAATCTCTGCAGCCAAATATGGGATCG 2635  
Qy 3060 ATCGCGACACCGCCAGAGACACATTTCTCTGCAATGACGACCTGACACCGA--A 3117  
Db 2636 GCCATTGAACAAGATGGATTGACGACGAGTTCTCCGGCCGTGTGGTGGAGAGGCTATTC 2695  
Qy 3118 GACACCGGGGTCTATTTTGTGCAAGCCTTACTTTCGGCTTCCCTGCTGTTTGTCTTATTGG 3177  
Db 2696 GGCTATGACTGGGCACACAGACATCGCTGCTCTGATGCGCGGTGTTCGGCTGTCA 2755  
Qy 3178 GGCAAGGAGACCCGGTCAACGCTCTCAGTGTCAAGCCCAACAGACGCCAGCGCCGGA 3237  
Db 2756 GCGAGGGGCGCCCGTCTTTTGTCAAGACCGACCTGTCCGGTGCCC---TGAATGAA 2812  
Qy 3238 CCAACCAACACCGCGCCCAACCATCGCTGCGAGCCCTGTCCCTGGCCCGCAGAGCGGCT 3297  
Db 2813 CTGCAGGACGAGGACGCGGCTATCGTGGCTGGCCACGACGGGCGTTCTTTCGCGCAGCT 2872  
Qy 3298 CGGCAGCGGGGGGGGCGAGTGACACAGAGGGGCTGGACTTCGCCCTGGATCCCAAA 3357  
Db 2873 GTGCTGACGTTGTCTACTGAAGGGAAGGACTGGCTGCTATTGGGCGAAGTGGCGGG 2932  
Qy 3358 CTCTGCTACTGCTGGATGGAATCCTCTTCACTATGTTGTCAATCTCACTGCTTGTTC 3417  
Db 2933 CAGGATCTCTGTCACTCACTTGTCTCTGCGAGAAAGTATCCATCATGGCTGATGCA 2992  
Qy 3418 CTGAGATGGAAGTTGACGAGAGCGAGAGCCCGCGGTACGACGAGGGCCAGACGAG 3477  
Db 2993 ATGCGGGGCTGCAATGATCGGCTTGTATCGGCTTACTTGCCTATTCGACCAACCAAGCGAAACAT 3052  
Qy 3478 CTCTATACGAGCTCAATCTAGGACGAGAGAGGATGATGATGTTTGGCAAGAGAGCT 3537  
Db 3053 CGC-----ATGACGAGACGCTACTCGGATGGAAGCCGGTCTTGTGATCAGGAT 3103  
Qy 3538 GCGCGGACCTCTGATGATGGGGGAAAGCCGAGAGGAAGAACCTTCAGAGAGGCTGTAC 3597  
Db 3104 GATCTGAGCAGAGACATCAGGG--GCTCGCGCAGCGGACTGTTCGCCAGGCTCAAGG 3161  
Qy 3598 AATGAATGCAAGAGATAAGATGCGGAGCGCTTACAGTGAATTTGGATGAAGAGGGAG 3657

Db 3162 CGCGCATCCCGCAGCGGCGAGGATCTCTGTCGTGACCCATGGCGATGCCTGCTTCCGAATA 3221  
Qy 3658 CGCGGAGGGCAAGG--GGCACGATGGCTTTTACCAGGGTCTCAGTACAGCCACCAAGG 3715  
Db 3222 TCATGGTGGAAATGGCGCTTTTCTGGATTCACTGAGTGTGGCGGCTGGGTGTGGCG 3281  
Qy 3716 ACACCTACGACGCGCTTTCACATGACGAGCCCTGCCCCCTCGCTAA-----CTCGACGGGGCC 3771  
Db 3282 ACCGCTATCAGGACATAGCGTTGGCTACCCCTGTATATTGCTGAAGAGCTTGGCGCGAAT 3341  
Qy 3772 CGCGATCCGATTAGTCCAATTTGTTAAAGACAGGATATCAGTGTCTCAGGCTCTAGTTT 3831  
Db 3342 GGGCTGACCGCTTCTCTGTGCTTTACGGTATCGCGCTCCCGATTCGCGAGCGCATCGCT 3401  
Qy 3832 TGACTCAACAATATCACCGAGCTGAAGCCTATAGAGTACGAGCCATAGATATAAATAAAGA 3891  
Db 3402 TCTATCGCTCTTTGACGAGTCTTCTGAGCGGACTCTGGGGTTCGATAAATAAATAAAGA 3461  
Qy 3892 TTTTATTTAGTCTCCAGAAAAGGGGGAATGAAGACCCCACTGTAGGTTTGGCAAGC 3951  
Db 3462 TTTTATTTAGTCTCCAGAAAAGGGGGAATGAAGACCCCACTGTAGGTTTGGCAAGC 3521  
Qy 3952 TAGCTTAAGTAAAGCCATTTTCAAGGCAATCG--AAAATACATACTGAGAAATAGAGAAT 4010  
Db 3522 TAGCTTAAGTAAAGCCATTTTGAAGGCTGGAAGAAATACATNACTGAGAAATAGAGAAT 3581  
Qy 4011 TCAGATCAAGGTTAGGAACAGA-GAGACAGCAGAAATATGGCCCAACAGGATATCTGTGG 4069  
Db 3582 TCAGATCAAGGTCAAGAACAGATGGAACAGCTGAATATGGGCCAAATATGGGCCAA 3641  
Qy 4070 TAAGCAGTTCTGCGCCC--GCTCAGGGCCAGAAACAGTTGGAAACAGGAAATATGGGCCAA 4128  
Db 3642 TAAGCAGTTCTGCGCCCGGCTCAGGGCCAAAGACAGATGGAACAGCTGAATATGGGCCAA 3701  
Qy 4129 ACAGGATATCTGTGTAGCAGTCTCTGCCCGGCTCAGGGCCAAAGACAGATGCTGCC 4188  
Db 3702 ACAGGATATCTGTGTAGCAGTCTCTGCCCGGCTCAGGGCCAAAGACAGATGCTGCC 3761  
Qy 4189 AGATCGGCTCCCGCCCTCAGCAGTCTTCTAGAGAACCATCAGATGTTTCCAGGGTCCCA 4248  
Db 3762 AGATCGGCTCCAGCCCTCAGCAGTCTTCTAGAGAACCATCAGATGTTTCCAGGGTCCCA 3821  
Qy 4249 AGGACCTGAAATGACCTGTGCTTATTGTAATCTAAACAATCAGTTTCGCTTCTCGCTTCT 4308  
Db 3822 AGGACCTGAAATGACCTGTGCTTATTGTAATCTAAACAATCAGTTTCGCTTCTCGCTTCT 3881  
Qy 4309 GTTCGCGGCTTCTGCTCCCGGAGCTCAATAAAGAGCCCAACCCCTCACTCGCGCG 4368  
Db 3882 GTTCGCGGCTTCTGCTCCCGGAGCTCAATAAAGAGCCCAACCCCTCACTCGCGCG 3941  
Qy 4369 CCAGTCTCCGATAGACTGCGTCCCGGGTACCCGTGTTCTCAATAAACCCCTCTTGCAG 4428  
Db 3942 CCAGTCTCCGATAGACTGAGTCCCGGGTACCCGTGTAATCAATAAACCCCTCTTGCAG 4001  
Qy 4429 TTGCATCCGACTCGTGTGCTGCTGCTTCTTGGGAGGGTCT-CTCTGAGTGAATGACTAC 4487  
Db 4002 TTGCATCCGACTGCTGCTGCTTCTTGGGAGGGTCTCTCTCTGAGTGAATGACTAC 4061  
Qy 4488 CCGTCAAGGGGCTTTCAGTCTTCTCCCACTACAGGCTCTCACTAATCTCTGATGT 4547  
Db 4062 CCGTCAAGGGGCTTTCATTTGGGGGCTCGTCCGGAGACCCCTGCGCCAGG 4121  
Qy 4548 GCGCAGGAGCTCCGTCAGCCCGGTTTGTGTTTATATAAATGCAAGAACAGATGTTTCCC 4607  
Db 4122 ACACACGACCAACCGGGAGGTAAAGT----- 4151  
Qy 4608 TTCAAGCCAGACTACATCTCGACTCTCGGCTTTTATAAAGAAATGTTGAAGGGCTCTG 4667  
Db 4152 -----GCTGCTCGCGGTTTCGGTGAATGACGCTGAAACCTCTG-AC 4193  
Qy 4668 ACTATCTGCCACAGACTTTTAAAGATTTTATGCTCTCGATGAGGGATTTAGTCAAT 4727  
Db 4194 ACATGCGAGCTCCCGGAGACGCTCAAGCTTGTCTGTAAGCGGATGCCGGAGCAGACAAG 4253



QY	4728	CTATCCTCGTCTATTTGCTGGCTTCTCCGTATTTTAAATTTCTAGTTTGCATCCCTTC	4787	5808	ACGACGCGAGTATGTAGGCGGTGCTACAGAGTTCTTTGAAGTGTGGCTTAACTCGCG	5867
Db	4254	CCCGTCAGGCGCGTTCAGCGGGTGTGGCG-----	4283	4990	AGCAGACGAGGTATGTAGGCGGTGCTACAGAGTTCTTTGAAGTGTGGCTTAACTCGCG	5049
QY	4788	CTGAGAGCAGCGGATGTCAGAGTAGTTAATCTCTGAGGGCAGGGCTTCTGTGAAGGT	4847	5868	TACACTAGACGACAGTATTTGGTATCTGCGCTCTGCTGAGCCAGTTACCTTCGGAAA	5927
Db	4284	-----	4283	5050	TACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAA	5109
QY	4848	TGCTGGGCTCAGTGTGAGATTTTGCCATAAAAAAGGGTCTTGCCCTCTGTACAGACAG	4907	5928	AGAGTTGGTAGTCTTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGTTTTTTTGT	5987
Db	4284	-----	4283	5110	AGAGTTGGTAGTCTTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGTTTTTTGT	5169
QY	4908	ATCGGAATCTAGAGTGCACTACTCAGAGTCCCGCGGTTCGGGGCTCTGATCTCAGGGCA	4967	5988	TGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCTTTTGATCTTTCT	6047
Db	4284	-----	4304	5170	TGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCTTTTGATCTTTCT	5229
QY	4968	TCCTTCCCTAGAGATCCTCTACGCCGACGCATCGTGGCGCGGTACCGAGCTCGAATCG	5027	6048	ACGGGGTCTGACGCTCAGTGGNAACGAAAACTCACGTTAAGGATTTTGGTCAATGAGATTA	6107
Db	4305	-----	4304	5230	ACGGGGTCTGACGCTCAGTGGNAACGAAAACTCACGTTAAGGATTTTGGTCAATGAGATTA	5289
QY	5028	TAATCATGTCTAGTGTGTTTCTGTGTGAATTTGTTATCCGCTCACAAATTCACACACAC	5087	6108	TCAAAAAGGATCTTCACCTTAGATCCTTTTAAATTAATAATGAAGTTTAAATCAATCTAA	6167
Db	4305	-----	4314	5290	TCAAAAAGGATCTTCACCTTAGATCCTTTTAAATTAATAATGAAGTTTAAATCAATCTAA	5349
QY	5088	ATACGAGCCGGAAGCATAAAGTGTAAAGCTCGGGTGCCTAATGAGTGAGCTAACTCACA	5147	6168	AGTATATATGAGTAAACTTTGGTCTGACAGTTTACCAATGCTTAATCAGTGAGGCACCTATC	6227
Db	4315	TAGCGATAGCGGATGTACTTGGCTTAATCTATCGGCATCAGACGAGATTGTACTGAGA	4374	5350	AGTATATATGAGTAAACTTTGGTCTGACAGTTTACCAATGCTTAATCAGTGAGGCACCTATC	5409
QY	5148	TTAATTTGCTTGGCTCACTGCGCGCTTTCCAGTGGGAAACCTGTGTCGACGTGCAT	5207	6228	TCAGGATCTGTCTATTTTGGTTCATCCATAGTTTGGCTGACTCCCGTCTGTGTAGATACT	6287
Db	4375	GT-----	4389	5410	TCAGGATCTGTCTATTTTGGTTCATCCATAGTTTGGCTGACTCCCGTCTGTGTAGATACT	5469
QY	5208	TAATGAATCGCCCAACGCGGGGAGAGCGGTTTGCCTATTTGGGGCTCTTCCGCTTCC	5267	6288	ACGATACGGGAGGCTTACCATCTGGCCCCAGTGTGCAATGATACCCGAGACCCACGC	6347
Db	4390	TGTGAATACCGACAGATGCGTAAAGGAGAAAAATACCGCATCAGCGCTCTTCCGCTTCC	4449	5470	ACGATACGGGAGGCTTACCATCTGGCCCCAGTGTGCAATGATACCCGAGACCCACGC	5529
QY	5268	TGCTCACTGACTCGCTCGGCTCGGCTCGGCTCGGCGAGCGGTATCAGTCACTCA	5327	6348	TCACCGGCTCCAGATTTTATCAGCAATAAACCCAGCCAGCGGAGGCGGAGCGAGAAGT	6407
Db	4450	TGCTCACTGACTCGCTCGGCTCGGCTCGGCTCGGCGAGCGGTATCAGTCACTCA	4509	5530	TCACCGGCTCCAGATTTTATCAGCAATAAACCCAGCCAGCGGAGGCGGAGCGAGAAGT	5589
QY	5328	AAGGGGTAAATACGGTTATCCACAGAAATCAGGGATTAACGAGGAAAGAACATGTGAGCA	5387	6408	GGTCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGGAAGCTAGATTA	6467
Db	4510	AAGGGGTAAATACGGTTATCCACAGAAATCAGGGATTAACGAGGAAAGAACATGTGAGCA	4569	5590	GGTCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGGAAGCTAGATTA	5649
QY	5388	AAGGCCAGCAAAAGCCAGAACCGTAAAGGCCGCTTGTGCTGGCGTTTTTCCATAGG	5447	6468	AGTAGTTCGCGAGTTAATAGTTTGGCAACGTTGTCATTTGCTTACAGGC-TCGTGTG	6526
Db	4570	AAAGGCCAGCAAAAGGCCAGAACCGTAAAGGCCGCTTGTGCTGGCGTTTTTCCATAGG	4629	5650	AGTAGTTCGCGAGTTAATAGTTTGGCAACGTTGTCATTTGCTTGCAGGCACTGTGTG	5709
QY	5448	CTCCGCCCTCTGACGAGCATCACAAAAATCGACGCTCAAGTCAAGTGGGCGAAACCCG	5507	6527	TCACGCTCGTTCGTTTGGTATGCTTCAATTCAGCTCCGCTTCCCAACGATCAAGGCGAGTT	6586
Db	4630	CTCCGCCCTCTGACGAGCATCACAAAAATCGACGCTCAAGTCAAGTGGGCGAAACCCG	4689	5710	TCACGCTCGTTCGTTTGGTATGCTTCAATTCAGCTCCGCTTCCCAACGATCAAGGCGAGTT	5769
QY	5508	ACAGGACTATAAGATACAGGGCTTTCCCGCTTGAAGCTCCCTCGTGGCTCTCCTGTT	5567	6587	ACATGATCCCCATGTTGCAAAAAAGGGTTAGTCTCTTCGGTCTCCGATCGTGTGC	6646
Db	4690	ACAGGACTATAAGATACAGGGCTTTCCCGCTTGAAGCTCCCTCGTGGCTCTCCTGTT	4749	5770	ACATGATCCCCATGTTGCAAAAAAGGGTTAGTCTCTTCGGTCTCCGATCGTGTGC	5829
QY	5568	CGACCCCTCGGCTTACCGGATACCTGTCCGCGCTTCTCCCTTCGGGAAAGCGTGGCGTT	5627	6647	AGAAAGTAGTTGGCGCGAGTGTATCACTCATGTTATGGCAGCACTGCATATTTCTCT	6706
Db	4750	CGACCCCTCGGCTTACCGGATACCTGTCCGCGCTTCTCCCTTCGGGAAAGCGTGGCGTT	4809	5830	AGAAAGTAGTTGGCGCGAGTGTATCACTCATGTTATGGCAGCACTGCATATTTCTCT	5889
QY	5628	TCTCATAGCTCAGGCTGTAGTATCTCAGTTCGGTGTAGTGTGCTTCCGCTCCAGCTGGGC	5687	6707	ACTGTTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCAAGTCATTC	6766
Db	4810	TCTCATAGCTCAGGCTGTAGTATCTCAGTTCGGTGTAGTGTGCTTCCGCTCCAGCTGGGC	4869	5890	ACTGTTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCAAGTCATTC	5949
QY	5688	TGTGTGACGAAACCCCGCTTCAGCCCGACCGCTGGCGCTTATCCGGTAACTATCGTCTT	5747	6767	TGAGAATAGTGTATCGGCGACCGAGTTGCTTTCGCCGGGTCAATACGGGATTAATACC	6826
Db	4870	TGTGTGCAAGAACCCCGCTTCAGCCCGACCGCTGGCGCTTATCCGGTAACTATCGTCTT	4929	5950	TGAGAATAGTGTATCGGCGACCGAGTTGCTTTCGCCGGGTCAACACGGGATTAATACC	6009
QY	5748	GAGTCCAAACCCGGTAAAGACAGCTTATCCGCACTGGGCAAGCCACTGGTAACAGGATT	5807	6827	GGGCCACATAGCAGAACTTTAAAGTGTCTCATTTGGAAACGTTCTTCGGGGGAAA	6886
Db	4930	GAGTCCAAACCCGGTAAAGACAGCTTATCCGCACTGGGCAAGCCACTGGTAACAGGATT	4989	6010	GGGCCACATAGCAGAACTTTAAAGTGTCTCATTTGGAAACGTTCTTCGGGGGAAA	6069
QY				6887	CTCTCAAGGATCTTACCGCTGTGTAGATCCAGTTCGATGTAAACCCACTCGTGCACCCAC	6946

```
Db 6070 CTCTCAAGATCTTACCGCTGTGTAGATCCAGTTCGATGTACCCACACTCGTGACCCCAAC 6129
Qy 6947 TGATCTTCAGCATCTTTTACTTTTACCTTCCAGCGTTTCTGGGTGAGCAAAACAGGAAGGCAA 7006
Db 6130 TGAATCTTCAGCATCTTTTACTTTTACCTTCCAGCGTTTCTGGGTGAGCAAAACAGGAAGGCAA 6189
Qy 7007 AATCGCGCAAAAGGGAATAGGCGGACACGGAATGTTGAATACTCATCTCTTCCTT 7066
Db 6190 AATCGCGCAAAAGGGAATAGGCGGACACGGAATGTTGAATACTCATCTCTTCCTT 6249
Qy 7067 TTTCAATATTATGAAGCATTTATCAGGGTTATGTTCTCATGAGCGGATACATATTGAA 7126
Db 6250 TTTCAATATTATGAAGCATTTATCAGGGTTATGTTCTCATGAGCGGATACATATTGAA 6309
Qy 7127 TGTATTTAGAAAAATAAACAATAGGGTTCCGCGCACATTTCCCGCAAAAGTGCACCT 7186
Db 6310 TGTATTTAGAAAAATAAACAATAGGGTTCCGCGCACATTTCCCGCAAAAGTGCACCT 6369
Qy 7187 GACGCTTAAGAAACCATTTATTATCATGACATTAACCTATAAAATAGCGGTATCACGAGG 7246
Db 6370 GACGCTTAAGAAACCATTTATTATCATGACATTAACCTATAAAATAGCGGTATCACGAGG 6429
Qy 7247 CCCTTTCGTCT 7257
Db 6430 CCCTTTCGTCT 6440
```

## RESULT 13

```
AX90482
ID AX90482 standard; DNA; 6522 BP.
XX
AC AX90482;
XX
DT 29-SEP-1999 (first entry)
XX
DE Plasmid retroviral vector pLUSN nucleotide sequence.
XX
KW Plasmid retroviral vector; expression system; immunogenic; gene therapy;
KW immune response; immunosuppression; gene delivery; therapeutic;
KW MHC-I autoimmune disease; tumour; ss.
XX
OS Synthetic.
XX
PN WO936562-A1.
XX
PD 22-JUL-1999.
XX
PF 13-JAN-1999; 98US-0071409P.
XX
PR 14-JAN-1999; 98US-0071409P.
XX
PA (HUMA-) HUMAN GENE THERAPY RES INST.
XX
PI Radosevich TJ, Link CJ;
XX
DR WPI; 1999-468988/39.
XX
PT Expression system containing therapeutic gene and an immunosuppressor
PT gene useful for treating an MHC-I autoimmune disease or killing tumor
PT cells.
XX
PS Claim 25; Page 137-140; 154pp; English.
XX
CC The present invention describes a nucleotide expression system for the
CC introduction of a therapeutic gene comprising: (i) a nucleotide sequence
CC encoding an immune suppression gene; (ii) a promoter; and (iii) a
CC transcription termination signal, where the system is able to inhibit,
CC evade or eliminate a recipient cell immune response to the therapeutic
CC gene when the gene is transformed into a recipient cell. AAX90481 to
CC AX90484 represent specifically claimed plasmid retroviral vector
CC nucleotide sequences from the present invention. The expression system
CC and vectors containing it can be used for gene therapy, for treating an
```

```
CC MHC-I autoimmune disease or for killing tumour cells. The expression
CC system contains an immunosuppressive gene which prevents host rejection
CC of the vector
XX
SQ Sequence 6522 BP; 1469 A; 1792 C; 1704 G; 1557 T; 0 U; 0 Other;
Query Match 40.4%; Score 3096; DB 2; Length 6522;
Best Local Similarity 70.8%; Pred. No. 0;
Matches 4803; Conservative 0; Mismatches 1465; Indels 518; Gaps 31;
Qy 549 TTTTGAAGAGCCACCCCTGTAGGTTTGGCAAGCTAGCTTAAAGTAACGCCATTTTCAAGG 608
Db 174 TTTTGAAGAGCCACCCCTGTAGGTTTGGCAAGCTAGCTTAAAGTAACGCCATTTTCAAGG 231
Qy 609 CATGG-AAAATACATACTGAGAAATAGAGAAGTTTCAGATCAAGGTTAGGAACAGAGAGAC 667
Db 232 CATGMAAAATACATACTGAGAAATAGAGAAGTTTCAGATCAAGGTTAGGAACAGAGAAAC 291
Qy 668 AGCAGATATGCGGCCAAACAGGATATCTGTGTAAAGCATTTCTGTCGCC- GTCAGGGCC 726
Db 292 AGCTGAATA-----CCAAACAGGATATCTGTGTAAAGCGTTTCTGCCCCGCTCAGGGCC 347
Qy 727 AAGAACAGTTTGGAAACAGAGAAAT-ATGGGCCAAACAGGATATCTGTGTAAAGCAGTTCT 785
Db 348 AAGAACAGATGAGACAGCTGAGTGATGGCCAAACAGGATATCTGTGTAAAGCAGTTCT 407
Qy 786 GCCCGGCTCAGGGCCAAAGAACAGATGGTCCCCAGATCGCGTCCCGCCCTCAGCAGTTTC 845
Db 408 GCCCGGCTCAGGGCCAAAGAACAGATGGTCCCCAGATCGCGTCCAGCCCTCAGCAGTTTC 467
Qy 846 TAGAGAACCATCAGATGTTTCCAGGGTCCCCAAAGGACCTG-AAATGACCTGTGCTTCA 904
Db 468 TAGTGAATCATCAGATGTTTCCAGGGTCCCCAAAGGACCTGAAAAATGACCTGTACCTTA 527
Qy 905 TTTCAACTTAACCAATCAGTTCTGCTTCTGCTTCTGTCGCGGCTTCTGCTCCCGAGCT 964
Db 528 TTTCAACTTAACCAATCAGTTCTGCTTCTGCTTCTGTCGCGGCTTCTGCTTCTCGAGCT 587
Qy 965 CAATAAAGAGCCCAACACCCCTCACTCGGCGCCAGTCTCTCGATAGACTGCGTCGCC 1024
Db 588 CAATAAAGAGCCCAACACCCCTCACTCGGCGCCAGTCTTCCGATAGACTGCGTCGCC 647
Qy 1025 CGGTACCCGTATTTCCCAATAAAGCCTTCTGCTGTTGTCATCCGAATCGTGAGCTCGCTG 1084
Db 648 CGGTACCCGTATTTCCCAATAAAGCCTTCTGCTGTTGTCATCCGAATCGTGCTCGCTG 707
Qy 1085 ATCCTTGGAGGGTCTCTCAGATTTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1144
Db 708 TTTCTTGGAGGGTCTCTCAGATTTGATGATGATGATGATGATGATGATGATGATGATGATG 767
Qy 1145 GGTTCACCCGAGATTTGGAGACCCCTGCCCAGGACCAACCGACCCCGCCCGGGAGGTA 1204
Db 768 GGCTGTCGGGATTTGGAGACCCCTGCCCAGGACCAACCGA-CCCACACCGGGAGGTA 826
Qy 1205 AGCTGGCCAGCAACTTATCTGTGTCGATTTCTAGTGTCTATGACTGATTTATG 1264
Db 827 AGCTGGCCAGCAACTTATCTGTGTCGATTTCTAGTGTCTATGTTGATGTTATG 886
Qy 1265 CGCTGCTGGTACTAGTACTAGTAACTAGTCTGATCTGTCGCGGACCCGCTGTTGAAT 1324
Db 887 CGCTGCTGGTACTAGTACTAGTAACTAGTCTGATCTGTCGCGGACCCGCTGTTGAAT 946
Qy 1325 GACGAGTTCCGAACACCCCGCCCAACCTCGGAGACGTCCTCAGGAGCTTCGGGGGCGGT 1384
Db 947 GACGAGTTCTGAACACCCCGCCCAACCTCGGAGACGTCCTCAGGAGCTTTGGGGGCGGT 1006
Qy 1385 TTTTGTGCGCCGACCTGAGTCTCTAAATCCCGATCGTTTAGGACTCTTTTGTGACACCCC 1444
Db 1007 TTTTGTGCGCCGACCTGAGGAAGGAGTCTGATGTTGAATCCGACCCCGTCA----- 1057
Qy 1445 CTTAGAGAGGGAATATGTTGTTCTGTTAGGAGAGAGAACCTATAAGTAGTTCCCGCTCC 1504
Db 1058 -----GGATATGTGTTCTTGGTAGGAGAGAGAACCTATAAGTAGTTCCCGCTCC 1107
```





```
QY 5813 AGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCCTAACTACGGCTACAC 5872
DB |||||||
QY 5073 AGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCCTAACTACGGCTACAC 5132
DB |||||||
QY 5873 TAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAAGCAGATTACCTTCGGAAAAAGAGT 5932
DB |||||||
QY 5133 TAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAAGCAGATTACCTTCGGAAAAAGAGT 5192
QY 5933 TGGTAGCTCTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGTTTTTTGTTGCAAA 5992
DB |||||||
QY 5193 TGGTAGCTCTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGTTTTTTGTTGCAAA 5252
QY 5993 GCAGCAGATTAGCGCAGAAAAAAGGATCTCAAGAAAGATCTCTTTGATCTTTTCTACGGG 6052
DB |||||||
QY 5253 GCAGCAGATTAGCGCAGAAAAAAGGATCTCAAGAAAGATCTCTTTGATCTTTTCTACGGG 5312
QY 6053 GTCTGACGTCAGTGGAACGAAACCTCACGTTAAGGGATTTGGTCATGAGATTAATCAAA 6112
DB |||||||
QY 5313 GTCTGACGTCAGTGGAACGAAACCTCACGTTAAGGGATTTGGTCATGAGATTAATCAAA 5372
QY 6113 AAGGATCTTCACTAGATCTTTTAAATTAATAATGAAGTTTTAAATCAATCTAAAGTAT 6172
DB |||||||
QY 5373 AAGGATCTTCACTAGATCTTTTAAATTAATAATGAAGTTTTAAATCAATCTAAAGTAT 5432
QY 6173 ATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTAGGCGACCTATCTCAGC 6232
DB |||||||
QY 5433 ATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTAGGCGACCTATCTCAGC 5492
QY 6233 GATCTGCTATTTGCTTCATCANTAGTTCGCTGACTCCCGCTCGTGTAGATAACTACGAT 6292
DB |||||||
QY 5493 GATCTGCTATTTGCTTCATCANTAGTTCGCTGACTCCCGCTCGTGTAGATAACTACGAT 5552
QY 6293 ACGGGAGGCTTACCATCTGGCCCGAGTCTGCAATGATACCGCGAGCCGACGCTCACC 6352
DB |||||||
QY 5553 ACGGGAGGCTTACCATCTGGCCCGAGTCTGCAATGATACCGCGAGCCGACGCTCACC 5612
QY 6353 GGCTCCAGATTTATCAGCAATAAACCAGCAGCGCGGAGCGCGAGCAAGTGGTCC 6412
DB |||||||
QY 5613 GGCTCCAGATTTATCAGCAATAAACCAGCAGCGCGGAGCGCGAGCAAGTGGTCC 5672
QY 6413 TGAACCTTTATCGGCTCCATCCAGTCTAATTAATTGTTCCGGGAAGCTAGAGTAAGTAG 6472
DB |||||||
QY 5673 TGAACCTTTATCGGCTCCATCCAGTCTAATTAATTGTTCCGGGAAGCTAGAGTAAGTAG 5732
QY 6473 TTGCGCAGTTAATAGTTTGGCAACGTTGTTGCCATTGCTACAGGC-TGCTGGTGTACAG 6531
DB |||||||
QY 5733 TTGCGCAGTTAATAGTTTGGCAACGTTGTTGCCATTGCTGAGGCAATCGTGGTGTACAG 5792
QY 6532 CTCGTCGTTTGGTATGGCTTTCAATTCAGCTCCGTTCCCAACGATCAAGGCGAGTTACATG 6591
DB |||||||
QY 5793 CTCGTCGTTTGGTATGGCTTTCAATTCAGCTCCGTTCCCAACGATCAAGGCGAGTTACATG 5852
QY 6592 ATCCCCCATGTTGTGCAAAAAAGCGGTAGCTCTTCGTTCCCTCCGATCGTTGTCAAGAG 6651
DB |||||||
QY 5853 ATCCCCCATGTTGTGCAAAAAAGCGGTAGCTCTTCGTTCCCTCCGATCGTTGTCAAGAG 5912
QY 6652 TAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATCAATTCCTTACTGT 6711
DB |||||||
QY 5913 TAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATCAATTCCTTACTGT 5972
QY 6712 CATGCCATCCGTAAGATGCTTTTCTGTGACTGTGTAGTACTCAACCAAGTCAATTCGAGA 6771
DB |||||||
QY 5973 CATGCCATCCGTAAGATGCTTTTCTGTGACTGTGTAGTACTCAACCAAGTCAATTCGAGA 6032
QY 6772 ATAGTGTATCGGCGCAGCGAGTTGCTCTTTCGCCGGCGTCAATACGGGATAAATACCGCGC 6831
DB |||||||
QY 6033 ATAGTGTATCGGCGCAGCGAGTTGCTCTTTCGCCGGCGTCAACACGGGATAAATACCGCGC 6092
QY 6832 ACATAGCAGAACTTTAAAGTGTCTCATCTTGGGAAACGTTCTTCGGGGCGGAAACTCTC 6891
DB |||||||
QY 6093 ACATAGCAGAACTTTAAAGTGTCTCATCTTGGGAAACGTTCTTCGGGGCGGAAACTCTC 6152
DB |||||||
```

```
QY 6892 AAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGCAACCACTGATC 6951
DB |||||||
QY 6153 AAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGCAACCACTGATC 6212
DB |||||||
QY 6952 TTACAGATCTTTTACTTTTCAACGAGGTTTCTGGGTGAGCAAAACAGGAGGCAAAATGC 7011
DB |||||||
QY 6213 TTACAGATCTTTTACTTTTCAACGAGGTTTCTGGGTGAGCAAAACAGGAGGCAAAATGC 6272
QY 7012 CCAAAAAAGGGAATAAAGGGCGACACGGAATGTTGAATACTCATACTCTCTCTCTTTTCA 7071
DB |||||||
QY 6273 CCAAAAAAGGGAATAAAGGGCGACACGGAATGTTGAATACTCATACTCTCTCTTTTCA 6332
QY 7072 ATATTATTGAAGCAATTTATCAGGGTTATGTCCTCATGAGCGGATACATATTTGAATGAT 7131
DB |||||||
QY 6333 ATATTATTGAAGCAATTTATCAGGGTTATGTCCTCATGAGCGGATACATATTTGAATGAT 6392
QY 7132 TTAGAAAAATAAACAAATAGGGTTCCGGGCAATTTCCCGGAAAGTGCACCTGACGT 7191
DB |||||||
QY 6393 TTAGAAAAATAAACAAATAGGGTTCCGGGCAATTTCCCGGAAAGTGCACCTGACGT 6452
QY 7192 CTAAGAAACCACTTATTATCATGACATTAACCTTATAAAATAGGCGTATCACGAGGCCCTT 7251
DB |||||||
QY 6453 CTAAGAAACCACTTATTATCATGACATTAACCTTATAAAATAGGCGTATCACGAGGCCCTT 6512
QY 7252 TCGTCT 7257
DB |||||||
QY 6513 TCGTCT 6518
DB |||||||

RESULT 14
AAZ34935
ID AAZ34935 standard; DNA; 6221 BP.
XX
AC AAZ34935;
XX
DT 28-FEB-2000 (first entry)
XX
DE Retrovirus vector.
XX
KW Interleukin-4 inducible epsilon promoter; immunoglobulin; IgE; antibody;
KW modulator; screening; human; allergy; therapy; retrovirus; vector;
KW green fluorescent protein; GFP; ss.
XX
OS Cytomegalovirus.
OS Aequorea victoria.
OS Unidentified.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT promoter 1..845
FT /tag= a
FT /note= "CMV promoter/R/US 5' LTR"
FT misc_feature 850..2100
FT /tag= c
FT /note= "extended psi region"
FT mutation replace(1322,G)
FT /tag= b
FT /note= "GAG ATG-ATC mutation"
FT misc_feature 2205..2723
FT /tag= d
FT /note= "ECMV IRES cloned from pCITE-4a"
FT CDS 2746..3465
FT /tag= e
FT /product= "green fluorescent protein"
FT LTR 3522..4115
FT /tag= f
FT /note= "3' LTR"
FT misc_feature 4122..62100
FT /tag= g
FT /note= "pGEM backbone (pUC origin, ampR)"
XX
XX WO9958663-A1.
XX
```



PD	18-NOV-1999.	QY	1437	GCACCCCTTAGAGGAGGATATGTTCTGTTAGGAGAGGAACTTAAACAGTTC	1496
XX					
PF	12-MAY-1999; 99WO-US010497.	Db	1140	GCACCCCTTAGAGGAGGATATGTTCTGTTAGGAGAGGAACTTAAACAGTTC	1199
XX					
PR	12-MAY-1998; 98US-00076624.	QY	1497	CCGCTCCGTCCTGAATTTTCTGTTTGGGACCGAACCGCGCGCTTGTTC	1556
XX					
PA	(RIGE-) RIGEL PHARM INC.	Db	1200	CCGCTCCGTCCTGAATTTTCTGTTTGGGACCGAACCGCGCGCTTGTTC	1259
XX					
PI	Ferrick DA, Swift SE, Armstrong R, Fox B;	QY	1557	TGCTGCAGCATCGTTCGTGTCTCTGTCTGATCTGTGTTTCTGTATTTGTCTGAAAT	1616
XX					
DR	WPI; 2000-062297/05.	Db	1260	TGCTGCAGCATCGTTCGTGTCTCTGTCTGATCTGTGTTTCTGTATTTGTCTGAAAT	1319
XX					
PT	Methods, cell lines and vectors for screening for modulators of	QY	1617	ATGGGCGCGGCTAGACTGTACCACTCCCTTAAGTTTGCCTTACCTTGGAAAGAT	1676
XX					
PT	immunoglobulin E synthesis, secretion and switch rearrangement.	Db	1320	ATCGGCGCGGCGGAGACTGTACCACTCCCTTAAAGTTTGCCTTACCTTGGAAAGAT	1379
XX					
PS	Disclosure; Fig 11A; 81pp; English.	QY	1677	GTCCAGCGGATCGCTCACACAGTCGGTAGATGTCAAGAAAGAGACGTTGGGTTACCTTC	1736
XX					
CC	This is the nucleotide sequence of a retrovirus vector preferred for use	Db	1380	GTCCAGCGGATCGCTCACACAGTCGGTAGATGTCAAGAAAGAGACGTTGGGTTACCTTC	1439
CC	in methods of the invention. The invention provides methods of screening	QY	1737	TGCTCTGCAGAAATGGCCAAACCTTTAAAGTCGGATGGCGCGGAGACGCTTTAAACCGA	1796
CC	for bioactive agents capable of inhibiting the human interleukin-4 (IL-4)	Db	1440	TGCTCTGCAGAAATGGCCAAACCTTTAAAGTCGGATGGCGCGGAGACGCTTTAAACCGA	1499
CC	inducible epsilon promoter (see AAZ34932). The method comprises combining	QY	1797	GACCTCATCACCAGGTTAAGATCAAGTCTTTTCACTGCGCGCGCATGGACACCCAGAC	1856
CC	a candidate bioactive agent with a cell comprising a fusion nucleic acid	Db	1500	GACCTCATCACCAGGTTAAGATCAAGTCTTTTCACTGCGCGCGCATGGACACCCAGAC	1559
CC	composed of the IL-4 inducible epsilon promoter and a reporter gene. The	QY	1857	CAGTCCCTTACATCGTCGACCTGGAGACCTTGGCTTTTGGACCCCTCCCTGGGTCAAG	1916
CC	promoter is then induced with IL-4 or IL-13, and the presence or absence	Db	1560	CAGTCCCTTACATCGTCGACCTGGAGACCTTGGCTTTTGGACCCCTCCCTGGGTCAAG	1619
CC	of the reporter gene is detected. Absence of the reporter gene indicates	QY	1917	CCCTTTGTACACCTTAAGCCTCCGCTCTCTCTCTCCATCCGCGCGCTCTCTCCCTTT	1976
CC	that the agent inhibits the promoter. Preferred embodiments use	Db	1620	CCCTTTGTACACCTTAAGCCTCCGCTCTCTCTCTCCATCCGCGCGCTCTCTCCCTTT	1679
CC	retroviral vectors to introduce the candidate bioactive agents. Also	QY	1977	GAACTCTCTCTGTTCCAGCCCGCTCGATCTCTCTCTTTATCCAGCGCTCACTCTCTCTA	2036
CC	provided are methods of screening for bioactive agents capable of	Db	1680	GAACTCTCTCTGTTCCAGCCCGCTCGATCTCTCTCTTTATCCAGCGCTCACTCTCTCTA	1739
CC	modulating IgE synthesis, secretion and switch rearrangement. These	QY	2037	GGCGCCCATATGGCCATATGAGATCTTATATGGGCGACCCCGCCCTTGTAACTTC	2096
CC	methods rely on reporter genes fused to IgE promoters, such as the IL-4	Db	1740	GGCGCCCATATGGCCATATGAGATCTTATATGGGCGACCCCGCCCTTGTAACTTC	1799
CC	inducible epsilon promoter that starts a cascade that ultimately results	QY	2097	CCTGACCTGACATGACAAAGAGTTACTAAACAGCCCTCTCTCCAAGCTCACTTACAGGT	2156
CC	in IgE production. The methods screen for upstream modulators of IgE	Db	1800	CCTGACCTGACATGACAAAGAGTTACTAAACAGCCCTCTCTCCAAGCTCACTTACAGGT	1859
CC	production to prevent the production of IgE and thus reduce or eliminate	QY	2157	-TCTACTTAGTCCAGCAGCAAGTCTGGAGACCTCTGGCGGAGCTTACCAAGAACTTC	2215
CC	the allergic response	Db	1860	CTCTACTTAGTCCAGCAGCAAGTCTGGAGACCTCTGGCGGAGCTTACCAAGAACTTC	1919
XX					
SQ	Sequence 6221 BP; 1471 A; 1738 C; 1548 G; 1464 T; 0 U; 0 Other;	QY	2216	GACCGACGGTGGTACCTCACCTTACCGAGTCCGCGGACACAGTGTGGGTCCGCGGAC	2275
	Query Match 40.4%; Score 3093; DB 3; Length 6221;	Db	2275	GACCGACGGTGGTACCTCACCTTACCGAGTCCGCGGAGCTTACCAAGAACTTC	1919
	Best Local Similarity 73.7%; Pred. No. 0;	QY	2276	CAGACTAAGAACCTTAGAACCTCGCTGGAAAGACCTTACACAGTCTCTGACCAACCTTC	2335
	Matches 4587; Conservative 0; Mismatches 960; Indels 678; Gaps 24;	Db	1980	CAGACTAAGAACCTTAGAACCTCGCTGGAAAGACCTTACACAGTCTCTGACCAACCTTC	2039
QY	958 CCGAGCTCAATAAAGAGCCACACACCTCACTCGGCGGCGGAGTCTCCGATAGACTG	QY	2336	ACCGCCCTCAAAGTAGACGGCATCGCATCTGGATACACGCGCGCCACCTGAAGGCTGCC	2395
Db	662 CAGAGCTCAATAAAGAGCCACACACCTCACTCGGCGGCGGAGTCTCCGATAGACTG	Db	2040	ACCGCCCTCAAAGTAGACGGCATCGC-GCTTGGATACACGCGCGCCACCTGAAGGCTGCC	2098
QY	1018 CGTGGCGGGTACCCGGTATCCCAATAAGCCTCTGCTGTTGATCCCGAATCGTGA	QY	2396	GACCGCGGGGTGACCATCTCTAGACTGCGATGGGATGGAGCTGTATCATCTCTTCT	2455
Db	722 AGTGGCGGGTACCCGGTATCCCAATAAGCCTCTGCTGTTGATCCCGAATCGTGA	Db	2099	GACCGCGGGGTGACCATCTCTAGACTGCG- - - - -GGATCTCGAGGG	2142
QY	1078 CTGCTGATCTTGGGAGGGTCTCTCTAGATTGATTGACTGCC-ACCTCGGGGGTCTTT	QY	2456	TGGTAGCAACAGCTACAGGTGTCCACTCCGACATCCAGCTGACCCAGAGCCACGACCC	2515
Db	781 CTGCTGTTCTTGGAGGGTCTCTCTAGATTGATTGACTGCC-ACCTCGGGGGTCTTT	Db	2143	ATCCACCAACCTAGGACCCCAATTAATTTGGAATTTCTTCAGCCCGGGGGATCCACTAGTT	2202
QY	1137 CATTTGGAGTTTCCACCGAGATTGGAGACCCCTGCGGAGGACACCGACCCCGGCC	QY	2516	TGAGCGCAGCGTGGGTGACAGAGTGACCATCACTCTGTAAAGCCAGTCAGGATGGGTA	2575
Db	841 CATTTGGGGTCTGTCGCGGATCGGAGACCCCTGCGGAGGACCCCGA-CCACACACC				
QY	1197 GGGAGTTAAGTGGCCAGCACTTATCTGCTCTGTCGATGTTGATGTTCTATGACTG				
Db	900 GGGAGTTAAGTGGCCAGCACTTATCTGCTCTGTCGATGTTGATGTTCTATGACTG				
QY	1257 ATTTTATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG				
Db	960 ATTTTATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG				
QY	1317 GTGGAACGTGACGAGTTCCGAAACACCGCGCGCAACCTCGGGAGAGCTCCAGGACCTTCG				
Db	1020 GTGGAACGTGACGAGTTCCGAAACACCGCGCGCAACCTCGGGAGAGCTCCAGGACCTTCG				
QY	1377 GGGGCGGTTTGTGGCCCGACCTGAGTCTCTTAAATCCCGATCGTTTAGGACTCTTTGGT				
Db	1080 GGGGCGGTTTGTGGCCCGACCTGAGTCCAAATAATCCCGATCGTTTGGACTCTTTGGT				

Db 2203 CTAGAGCGAATTAAATTCGGTTATTATTCCACCA-----TA 2237  
Qy 2576 CTTCTGTAGCTTGGTACCAGCAGAGCCAGGTAAAGGCTCCAAAGCTGCTGATCTACTGGA 2635  
Db 2238 TTGCGGCTCTTTTGGCAGATGTAGGGCCCGGAACCTGGGCCCTGTCTTCTTGAGAGCAATT 2297  
Qy 2636 CATCCACCGCGCACACTGTGTGCCAAGCAGATTTACAGCGGTAGCGGTAGCGGTACCGACT 2695  
Db 2298 CCTAGGGGTCTTTCCCTCTCGCRAAAGGAATGCAAGGTCTGTGAATGTGCTGAAGGAA 2357  
Qy 2696 TCACCTTCAACATCAGCAGCGCTCCAGCCAGAGGACATCGCCACCTACTACTCTGCGAGCAAT 2755  
Db 2358 GCAGTTCTCTGGAAGCTTTCTGAAGACAAACACGTCTGTAGCGACCCCT----- 2408  
Qy 2756 ATAGCCTCTATCGGTGTTGCGCCAGGACCAAGGTGGAATCAACAGAGGTGGCTCAG 2815  
Db 2409 -----TGAGGCGAGCGGAACCCCCCACTGGCGACAGGTGCTCTGCG 2451  
Qy 2816 GATCGGGTGGATCCGCTCTGTGGCTCAGGATCGGAGGTCCAACTGTGGAGAGCGGTG 2875  
Db 2452 GCCAAAGCCA CGTGTATAGATACACTTGCAAGGCGGCACNACCCAGTGCCAGTTG 2511  
Qy 2876 GAGGTGTTGCAACCTGCGCGGTCCCTGCGCTGCTGCTCGCATCTGGCTTGCAATT 2935  
Db 2512 TGAGTTGGATAGTTGTGGAAGAGTCAAAATGGCTCTCTCAAGCGTATTCAACAAGGGGC 2571  
Qy 2936 TCACCACATATTGGATAGTTGGGTGAGACAGCACCTGGAAAAGTCTTGATGGATTG 2995  
Db 2572 TGAAGGATGCCAAGAGGTACCCCATTTGTATGGGATCTGATCTGGGGCTCTGGTGACAT 2631  
Qy 2996 GAGAAATTCATCAGATAGCAGTACGATTAACTATGCGCGCTCTTAAAGGATAGATTTA 3055  
Db 2632 GCTTTACATGTGTTAGTCGAGTTTAAAAAGCTCTAGGCCCCCGAACCA CGGGACGT 2691  
Qy 3056 CAATATCGGAGACAAACGCAAGAACACATTTGTTCTGCAATAGGACGCTGAGACCG 3115  
Db 2692 GGTTTTCTTTGAAAACACGATGATAATG-----GGGGATCCACCG 2736  
Qy 3116 AAGACACCGGGTCTATTTTGTGCAAGCTTTACTTCGGCTTCCCTGTTGCTTATT 3175  
Db 2737 TCGCCACCATGTGTGACAAAGGCGGAGAGCTGTTCACCGGGTGG--TGCCCATCTGCT 2794  
Qy 3176 GGGGCCAAGGACCCCGGTCACTGCTCCAGTGTCTAAGCCACACGACGCGCAGCGCCGC 3235  
Db 2795 CGAGCTGGACGGCAGTAAACGGCCCAAGTTACGCTGTGCGGCGAGGGGAGGCGCA 2854  
Qy 3236 GACCACCAACACGGCGCCACCATCGGTGCGAGCCCTGCTGCTGCGCCCGCAGAGGGG 3295  
Db 2855 TGCACCTACGGCA-AGCTGACCTGAAGTTTCATCTGCACACCGCGCAAGCTGCGCGTGC 2913  
Qy 3296 CTGGCCAGCGGGGGGGGCGAGTGACACAGAGGGGCTGAGACTTCGGCCCTGGATCCCA 3355  
Db 2914 CTTGGCCCACTCTGTTGACACCTGACCTGACCTGAGCTGAGTGTCTTTCAGCGGTACCCG 2973  
Qy 3356 AACTGTGCTACCTGCTGGATGGAATCCTCTTCACTATGCTGTCTTCTCACTGCTGCTGT 3415  
Db 2974 ACCAATGAGACGACGA-----CTTCTTAAGTCCGCCATCGCCCAAGGGTACG 3024  
Qy 3416 TCCGTAGAGTGAAGTTTCAGCAGAGCGCAGAGCCCGCGGTACCAAGAGGGCCAGAAC 3475  
Db 3025 TCCAGGAGCGCACCATCTTCTTCAAGGACGAGCGCAACTACAAAGACCCGCGCGAGTGCA 3084  
Qy 3476 AGCTCTATACGAGCTCAATCTAGGACGAAGAGAGTACGATGTTTTGGACAGAGAC 3535  
Db 3085 AGTTCGAGGGCGACACCTGTTGAACCGCATCGAGCTGAAGGGCATTCGATTTCAAGGAGG 3144  
Qy 3536 GTGGCCGGACCTGTAGATGGGGGAAAGCCGAGAGGAGAACCTCAGG---AAGGCC 3592  
Db 3145 ACGGCAACATCTGGGGGCAAGCTGGAGTACACTTACACAGCCCAACAGCTCTATATCA 3204  
Qy 3593 TGTACAAATGACATGCAAGAAGATTAAGATGGCGAGGCCCTACAGTGAATTTGGGATGAAAG 3652

3205 TGGCGCAAGCAGAGAAAGCGGCATCAAGTGAATCTCAAGATCCGCCACAACATCGAGG 3264  
3653 GCGAGCGCGGAGGCGCAAGGGGCACGATGGCTTTTACAGGGTCTCAGTACAGCCACCA 3712  
3265 ACGGAGCGGTGAGCTCGCCGACCACTACAGCAGAACACCCCCATCGGCGAGCGCCCG 3324  
3713 AGGACACCTAACAACCGCTTCAATGCGAGGCCCTCGCCCTCGCTAACTCTGACGGCGCG 3772  
3325 TGCTGCTGCCCAACCAACCACTACCTGAGCACCCGCTCGCCCTGAGCAAGAACCCCAACG 3384  
3773 CGGATCCCGATTAGTCCCAATTTGTTAAAGACAGGATATCAGTGTGCTCAGGCTCTAGTTT 3832  
3385 AGAAGCGCGATCA-----CATGCTCTGCTGGAGTTTCGTGACCGCCCGCGGAT 3433  
3833 GACTCAACAATATCACCAGCTGAAGCTATAGAGTACGAGCCATAGATAAAATAAAGAT 3892  
3434 CACTCTCGCATGAGAGAGCTGTACAAGTAAGCGCGCTCGACGATAAAATAAAGAT 3493  
3893 TTTATTTAGTCTCCAGAAAAGGGGGGAATGAAGACCCACCTGTAGGTTTGGCAAGCT 3952  
3494 TTTATTTAGTCTCCAGAAAAGGGGGGAATGAAGACCCACCTGTAGTGTGGCAAGCT 3553  
3953 AGCTTAAGTAAGCCATTTTGAAGGCACTGG-AAAATACATAACTGAGNATAGAGAGTT 4011  
3554 AGCTTAAGTAAGCCATTTTGAAGGCACTGGAAAATAATACATACTGAGNATAGAGAGTT 3613  
4012 CAGATCAAGTTTGAAGACAGA-GAGACAGCAGAAATATGGGCCAAACAGGATATCTGTGT 4070  
3614 CAGATCAAGTCTAGGACAGATGGAACAGCTGNAATATGGGCCAAACAGGATATCTGTGT 3673  
4071 AAGCAGTTTCTGCCCC-GTCCAGGGCCAAAGAACAGTTTGAACAGAGGAATATGGGCCAAA 4129  
3674 AAGCAGTTTCTGCCCCGCTCAGGCGCCAAAGAACAGATGGAACAGCTGAATATGGGCCAAA 3733  
4130 CAGGATATCTGTGTTAAGAGTTCTGCGCCCGGCTCAGGGCCAAAGAACAGATGTTCCCA 4189  
3734 CAGGATATCTGTGTTAAGCAGTTCTGCCCCGCTCAGGGCCAAAGAACAGATGTTCCCA 3793  
4190 GATCGGTCCCGCTCAGCAGCTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCCCAA 4249  
3794 GATCGGTCCAGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCCCAA 3853  
4250 GGACCTGAAATGACCTGTGCTTATTTGAATCAACCAATCAGTTTCTGCTTCTCGTTCTG 4309  
3854 GGACCTGAAATGACCTGTGCTTATTTGAATCAACCAATCAGTTTCTGCTTCTCGTTCTG 3913  
4310 TTCGCGCGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGGCGC 4369  
3914 TTCGCGCGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGGCGC 3973  
4370 CAGTCTCCGATAGACTGCTCGCGCGGTACCGGTGTTCTCAATAAACCCCTCTTGCAGT 4429  
3974 CAGTCTCCGATGAGTCTGCTCGCGCGGTACCGGTGATTCANATAAACCCCTCTTGCAGT 4033  
4430 TGCATCCGACTCGTGTCTGCTGTTCTTGGAGGGTCT-CTCTGAGTGATTGACTACC 4488  
4034 TGCATCCGACTGTTGTGTCTGCTGTTCTTGGAGGGTCTCTCTGAGTGATTGACTACC 4093  
4489 CTTGAGCGGGGCTTCTTCACTTCTCCCACTACAGGCTCTCACTAACTTCTGATG 4548  
4094 CTTGAGCGGGGCTTCTTCACT----- 4114  
4549 CCGCAGGAGCTCGGTGAGCCCGGTTTGTGTTTATAAATAAAGAACAGTGTTCCT 4608  
4115 ----- 4114  
4609 TCAAGCCAGACTACATCTGACTCTCGGCTTTTATAAAGATGTTTGAAGGCTCTGTGGA 4668  
4115 ----- 4114  
4669 CTATCTGCCACAGCACTTTTAAAGATTTTATGCTCTCTGGATGAGGGATTTAGTCAATC 4728  
4115 ----- 4114

QY 4729 TATCCTCGTCTATTTTGTGGCTTCTCCGTATTTTAAATTTCTAGTTTGCATCCCTTCC 4788  
DB 4115 ----- 4114  
QY 4789 TGAGAGCACGGGATTTGCAGAGTAGTTAATCTCTGAGGCGAGGCTTCTGTGAAAAGGTT 4848  
DB 4115 -----TCGACTTTGGTCTCGTGCCTTGGGA 4142  
QY 4849 GCCTGGGCTCAGTGTGAGATTGTCATATAAAGGGGTCCTGCCCTCTGTACAGACAGA 4908  
DB 4143 GGGTCTCTCTGAGTGAATGACTACCGCTCAGCGGGGCTCTTCAATGC----- 4191  
QY 4909 TCGGAATCTAGAGTGCACTACTCAGAGTCCC CGCGGTTCCGGGGCTCTGATCTCAGGGCAT 4968  
DB 4192 ----- 4191  
QY 4969 CTTTGCCTAGAGATCCTCTACGCCGAGCATCGTGGCGGGTACCGAGCTCGAATTCGT 5028  
DB 4192 ----- 4191  
QY 5029 AATCATGTGTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAATTCACACACA 5088  
DB 4192 -----AGCATGATCAAAATTAATTTGGTTTTTTTCTTAGTATTACATTA 4239  
QY 5089 TACGAGCCGGAAGCATATAAGTGTAAAGCTCGGGTGCCTAATGAGTGAGCTAACTCAAT 5148  
DB 4240 AA----- 4241  
QY 5149 TAAATTCGTTGGCTCACTGCCCGCTTTCAGTCGGGAAACCTGTCGTGGCAGCTGCATT 5208  
DB 4242 -----TGCCCATAGTTGCATT 4257  
QY 5209 AATGAATCGGCCAACCGCGGGAGAGCGGTTTTCGTTATTTGGGCGCTCTTCCGCTTCT 5268  
DB 4258 AATGAATCGGCCAACCGCGGGAGAGCGGTTTTCGTTATTTGGGCTCTTCCGCTTCT 4316  
QY 5269 CGCTCACTGA CTGCTGCGCTCGGTTCGTTTCGCTGCGCGAGCGGTATFACGCTCACTCAA 5328  
DB 4317 CGCTCACTGA CTGCTGCGCTCGGTTCGTTTCGCTGCGCGAGCGGTATFACGCTCACTCAA 4376  
QY 5329 AGGCGGTAAATACGGTTATCCAGATCAGGGGATTAAGCGAGGAAGACATGTAGCAA 5388  
DB 4377 AGGCGGTAAATACGGTTATFCCAAGATCAGGGGATAACCGAGGAAGAACATGTAGCAA 4436  
QY 5389 AAGGCCAGCAAAAGGCCAGGAACCGTAAAGGCGCGTTCGTCGCGTTTTTCCATAGGC 5448  
DB 4437 AAGGCCAGCAAAAGGCCAGGAACCGTAAAGGCGCGTTCGTCGCGTTTTTCCATAGGC 4496  
QY 5449 TCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGA 5508  
DB 4497 TCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGA 4556  
QY 5509 CAGGACTATAAGATACAGCGGTTTCCCTGGAAGCTCCCTCGTGGCTCTCTGTTC 5568  
DB 4557 CAGGACTATAAGATACAGCGGTTTCCCTGGAAGCTCCCTCGTGGCTCTCTGTTC 4616  
QY 5569 CGACCTCGCCCTTACCGGATACCTGTCGCGCTTCTCCCTTCGSGAAGCGTGGCTTT 5628  
DB 4617 CGACCTCGCCCTTACCGGATACCTGTCGCGCTTCTCCCTTCGSGAAGCGTGGCTTT 4676  
QY 5629 CTCATAGCTCAGCTGTAGTATCTCAGTTTCGTTGAGGTGTTTCGCTCCAAAGCTGGGCT 5688  
DB 4677 CTCATAGCTCAGCTGTAGTATCTCAGTTTCGTTGAGGTGTTTCGCTCCAAAGCTGGGCT 4736  
QY 5689 GTGTGCAAGAACCCCGCTTACGCGGACCGCTGCGCTTATCCGGTAACTATCGTCTTG 5748  
DB 4737 GTGTGCAAGAACCCCGCTTACGCGGACCGCTGCGCTTATCCGGTAACTATCGTCTTG 4796  
QY 5749 AGTCCAAACCCGCTAAGACACCACTATTCGACCTGCGCAGCGCACCTGTTAAACAGGATTA 5808  
DB 4797 AGTCCAAACCCGCTAAGACACCACTATTCGACCTGCGCAGCGCACCTGTTAAACAGGATTA 4856

QY 5809 GCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTCAAGTGTGGCTAACTACGGCT 5868  
DB 4857 GCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTCAAGTGTGGCTAACTACGGCT 4916  
QY 5869 ACACCTAGAGACAGTATTTGGTATCTCGCTCTGCTCAAGCCAGTATACCTTCGAAAAA 5928  
DB 4917 ACACCTAGAGACAGTATTTGGTATCTCGCTCTGCTCAAGCCAGTATACCTTCGAAAAA 4976  
QY 5929 GAGTTGGTAGCTCTTTGATCCGGCAAAACAAACACCGCTGTGTAGGGTGGTTTTTTTGT 5988  
DB 4977 GAGTTGGTAGCTCTTTGATCCGGCAAAACAAACACCGCTGTGTAGGGTGGTTTTTTTGT 5036  
QY 5989 GCAAGCAGCAGATTACGCGCAGAAAAAAGAGTCTCAAGAAAGATCCTTTGATCTTTCTA 6048  
DB 5037 GCAAGCAGCAGATTACGCGCAGAAAAAAGAGTCTCAAGAAAGATCCTTTGATCTTTCTA 5096  
QY 6049 CGGGTCTGACGCTCAGTGGAAACCAACTCAAGTAAAGGATTTTGGTCAATGAGATTAT 6108  
DB 5097 CGGGTCTGACGCTCAGTGGAAACCAACTCAAGTAAAGGATTTTGGTCAATGAGATTAT 5156  
QY 6109 CAAAAAGGATCTTCACTAGATCCTTTTAAATTAATAAATGAAGTTT---TAAATCAATCT 6165  
DB 5157 CAAAAAGGATCTTCACTAGATCCTTTTAAATTAATAAATGAAGTTTGGCGCAATCAATCT 5216  
QY 6166 AAAGTATATAGATTAACCTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGACCTA 6225  
DB 5217 AAAGTATATAGATTAACCTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGACCTA 5276  
QY 6226 TCTCAGCGATCTGCTATTTTGGTTCATCCATAGTTGCTGACTCCCGTCTGTAGATAA 6285  
DB 5277 TCTCAGCGATCTGCTATTTTGGTTCATCCATAGTTGCTGACTCCCGTCTGTAGATAA 5336  
QY 6286 CTACGATACGGGAGGCTTACCATCTGCGCCAGTGTGCAATGATACCGCGAGACCCAC 6345  
DB 5337 CTACGATACGGGAGGCTTACCATCTGCGCCAGTGTGCAATGATACCGCGAGACCCAC 5396  
QY 6346 GCTCACCGCTCCAGATTTATFAGCAATAAACACGAGCGGAGGCGGAGGCGAGAA 6405  
DB 5397 GCTCACCGCTCCAGATTTATFAGCAATAAACACGAGCGGAGGCGGAGGCGAGAA 5456  
QY 6406 GTGCTCTGCAACTTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGGAAGCTAGAG 6465  
DB 5457 GTGCTCTGCAACTTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGGAAGCTAGAG 5516  
QY 6466 TAAGTAGTTCGCGAGTTAATAGTTTGCACAACGTTTGGCATTGCTACAGGC--TCGTGG 6524  
DB 5517 TAAGTAGTTCGCGAGTTAATAGTTTGCACAACGTTTGGCATTGCTACAGGCATCTGTCG 5576  
QY 6525 TGTCAOGCTCGTCTGTTGGTATGSCCTTCATTCAGCTCCGTTCCCAACGATCAAGGCGAG 6584  
DB 5577 TGTCAOGCTCGTCTGTTGGTATGSCCTTCATTCAGCTCCGTTCCCAACGATCAAGGCGAG 5636  
QY 6585 TTACATGATCCCCATGTTGTGCAAAAAGGGTTAGCTCCTTCGGTCTCTCGATCTGTTG 6644  
DB 5637 TTACATGATCCCCATGTTGTGCAAAAAGGGTTAGCTCCTTCGGTCTCTCGATCTGTTG 5696  
QY 6645 TCAGAGTAAAGTTGGCGCAGTGTATCACTCATGTTTATGGCAGCACTGATTAATCTC 6704  
DB 5697 TCAGAGTAAAGTTGGCGCAGTGTATCACTCATGTTTATGGCAGCACTGATTAATCTC 5756  
QY 6705 TTACTGTCAATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGTGACTCAACCAAGTCA 6764  
DB 5757 TTACTGTCAATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGTGACTCAACCAAGTCA 5816  
QY 6765 TCTGAGAAATAGTATGCGGGGACCGAGTTGCTCTTCCCGCGGCTCAATACGGGATATA 6824  
DB 5817 TCTGAGAAATAGTATGCGGGGACCGAGTTGCTCTTCCCGCGGCTCAACACGGGATATA 5876  
QY 6825 CCGCGCACATAGCAGAACTTTTAAAGTGTCTCATATTGGAACAGTTCTTTCGGGGCGAA 6884  
DB 5877 CCGCGCACATAGCAGAACTTTTAAAGTGTCTCATATTGGAACAGTTCTTTCGGGGCGAA 5936  
QY 6885 AACTCTCAGGATCTTACCCTGTTGAGATCCAGTTTCGATGTAAACCCACTCGTGCAACCA 6944

Db 5937 AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATTAACCCACTCGTCGACCCA 5996  
 Qy 6945 ACTGATCTTCAGCATCTTTTACTTTTACCAGGCTTTCTGGGTGAGCAAAAACAGGAGGC 7004  
 Db 5997 ACTGATCTTCAGCATCTTTTACTTTTACCAGGCTTTCTGGGTGAGCAAAAACAGGAGGC 6056  
 Qy 7005 AAAATGCCCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCC 7064  
 Db 6057 AAAATGCCCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCC 6116  
 Qy 7065 TTTTCAATATTATTGAACATTTATFACGGGTTATTGTCTCATGAGCGGATACATATTG 7124  
 Db 6117 TTTTCAATATTATTGAACATTTATFACGGGTTATTGTCTCATGAGCGGATACATATTG 6176  
 Qy 7125 AATGATTTAGAAAAATAAACAATAGGGGTTCCGGCGACATTC 7169  
 Db 6177 AATGATTTAGAAAAATAAACAATAGGGGTTCCGGCGACATTC 6221

## RESULT 15

AAF30944  
 ID AAF30944 standard; DNA; 6221 BP.

AC AAF30944;

DT 11-SEP-2003 (revised)

DT 23-JUL-2001 (first entry)

XX Vector used in invention.

XX Interleukin-4 inducible epsilon promoter; human; IgE; antibody;

KW immunoglobulin E; allergy; therapy; switch rearrangement; vector; CMV;

KW green fluorescent protein; ds.

OS Cytomegalovirus.

OS unidentified.

OS Chimeric.

FH Key Location/Qualifiers

FT LTR 1..845

FT /tag= a

FT /note= "5' LTR"

FT promoter 1..845

FT /tag= b

FT /note= "CMV promoter"

FT misc\_feature 850..2100

FT /tag= d

FT /note= "extended psi region"

FT mutation replace(1322,G)

FT /tag= c

FT misc\_feature 2146..2173

FT /tag= e

FT /note= "2 Bstx1 peptide cloning sites"

FT misc\_feature 2205..2723

FT /tag= f

FT /note= "2CMV IRES cloned as EcoRI/MscI fragment from pCITE-4a"

FT CDS 2745..3464

FT /tag= g

FT /product= "green fluorescent protein"

FT LTR 3522..4115

FT /tag= h

FT /note= "3' LTR"

FT misc\_feature 4122..6210

FT /tag= i

FT /note= "pGEM backbone (pUC origin, ampr)"

XX WO200134806-A2.

XX 17-MAY-2001.

XX 13-NOV-2000; 2000WO-US031232.

PF

XX 12-NOV-1999; 99US-0165189P.  
 PR (RIGE-) RIGEL PHARM INC.  
 XX Kinsella TW;  
 PI WPI; 2001-335931/35.  
 XX Screening for agents capable of inhibiting a promoter, especially  
 PT interleukin-4 inducible epsilon promoter involved in immunoglobulin E  
 PT production, by using diphtheria toxin constructs.  
 XX Disclosure; Fig 11A-1-11A-3; 80pp; English.  
 XX The present sequence is that of a vector preferred for use in methods of  
 CC the invention. These methods utilise diphtheria toxin for screening  
 CC purposes, especially for identifying modulators of IgE synthesis,  
 CC secretion and switch rearrangement. A claimed method of screening for  
 CC bioactive agents capable of inhibiting the IL-4 inducible epsilon  
 CC promoter, which is involved in IgE switching, comprises: combining a  
 CC candidate bioactive agent and a cell that does not endogenously express  
 CC heparin-binding epidermal growth factor-like growth factor (HBEGF) and  
 CC which comprises a fusion nucleic acid encoding HBEGF; inducing the promoter  
 CC epsilon promoter and a nucleic acid encoding HBEGF; inducing the promoter  
 CC with IL-4; adding diphtheria toxin to the cell; and determining whether  
 CC the cell is dead. Compositions comprising a test vector and a reporter  
 CC vector that includes a reporter gene such as green fluorescent proteins  
 CC are provided. Inhibitors of IgE synthesis can be identified that prevent  
 CC the production of IgE and reduce or eliminate an allergic response.  
 CC (Updated on 11-SEP-2003 to standardise OS field)  
 XX  
 SQ Sequence 6221 BP; 1471 A; 1738 C; 1548 G; 1464 T; 0 U; 0 Other;  
 Query Match 40.4%; Score 3093; DB 4; Length 6221;  
 Best Local Similarity 73.7%; Pred. No. 0;  
 Matches 4587; Conservative 0; Mismatches 960; Indels 678; Gaps 24;  
 Qy 958 CCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGGCGCCAGTCTCCGATGACTG 1017  
 Db 662 CAGAGCTCAATAAAGAGCCCAACCCCTCACTCGGGCGCCAGTCTCCGATGACTG 721  
 Qy 1018 CGTCGCCGGGTACCCGTATTCCTCAATAAAGCTCTTGTGTGTGCATCCGAATCGTGA 1077  
 Db 722 AGTCGCCGGGTACCCGTATTCCTCAATAAAGCTCTTGTGTGTGCATCCGAATCGTGA 780  
 Qy 1078 CTCGCTGATCTCTGGGAGGGTCTCTCAGATTGATGACTGCC-ACCTCGGGGGTCTTT 1136  
 Db 781 CTCGCTGTTCTTGGGAGGGTCTCTCAGATTGATGACTGCCCTCAGCGGGGGTCTTT 840  
 Qy 1137 CATTTGGAGGTTCCACCGAGATTTTGGAGACCCCTCCAGGAGCACCAGACCCCGCC 1196  
 Db 841 CATTTGGGGGCTCGTCCGGGATCGGGAGACCCCTCCAGGAGCACCAG-CCACACC 899  
 Qy 1197 GGGAGGTAAGCTGGCCAGCACTTATCTGTGTCTGTCCGATTGTCTAGTGTCTATGACTG 1256  
 Db 900 GGGAGGTAAGCTGGCCAGCACTTATCTGTGTCTGTCCGATTGTCTAGTGTCTATGACTG 959  
 Qy 1257 ATTTATGCGCTCGTGGTACTAGTAACTAGCTCTGTATCTGGCGGACCCGCTG 1316  
 Db 960 ATTTATGCGCTCGTGGTACTAGTAACTAGCTCTGTATCTGGCGGACCCGCTG 1019  
 Qy 1317 GTGGAACGACGAGTTCGGAACACCCGCCCAACCTCCGGGAGAGCTCCACGGACTTCG 1376  
 Db 1020 GTGGAACGACGAGTTCGGAACACCCGCCCAACCTCCGGGAGAGCTCCACGGACTTCG 1079  
 Qy 1377 GGGGCGGTTTTGTGGCCCGACCTGAGTCTTAAATCCCGATCGTTTAGACTCTTTGGT 1436  
 Db 1080 GGGGCGGTTTTGTGGCCCGACCTGAGTCTTAAATCCCGATCGTTTAGACTCTTTGGT 1139  
 Qy 1437 GCACCCCTTAGAGAGGGATATGTGGTCTGTGGTAGGAGACGAGAACCTAAACAGTTTC 1496  
 Db 1140 GCACCCCTTAGAGAGGGATATGTGGTCTGTGGTAGGAGACGAGAACCTAAACAGTTTC 1199







QY 5869 ACACATAGAAAGCAGATATTTGGTATCTCGGCTCTGCTGAGCGAGTTACCTTCGGAAAAA 5928  
Db |||||  
QY 4917 ACACATAGAAAGCAGATATTTGGTATCTCGGCTCTGCTGAGCGAGTTACCTTCGGAAAAA 4976  
Db |||||  
QY 5929 GAGTTGGTAGCTCTTGATCCGCGCAAAACAAACACCGCTGGTAGCGGTGTTTTTGTGTT 5988  
Db |||||  
QY 4977 GAGTTGGTAGCTCTTGATCCGCGCAAAACAAACACCGCTGGTAGCGGTGTTTTTGTGTT 5036  
Db |||||  
QY 5989 GCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAAGATCCCTTTGATCTTTTCTA 6048  
Db |||||  
QY 5037 GCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAAGATCCCTTTGATCTTTTCTA 5096  
Db |||||  
QY 6049 CGGGCTCTGACGCTCAGTGGAACGAAACTCACGTTAAGGATTTTGGTCAAGATAT 6108  
Db |||||  
QY 5097 CGGGCTCTGACGCTCAGTGGAACGAAACTCACGTTAAGGATTTTGGTCAAGATAT 5156  
Db |||||  
QY 6109 CAAAAAGGATCTTACACCTAGATCTTTTAAATTAATAATGAAGTTT---TAAATCAATCT 6165  
Db |||||  
QY 5157 CAAAAAGGATCTTACACCTAGATCTTTTAAATTAATAATGAAGTTTGGCAAAATCAATCT 5216  
Db |||||  
QY 6166 AAAGTATATATCAGTAAACTTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTA 6225  
Db |||||  
QY 5217 AAAGTATATATGAGTAAACTTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTA 5276  
Db |||||  
QY 6226 TCTCAGCGATCTGTCTATTTCTGTTCAATCCATAGTTGCTGACTCCCGTCTGTAGATAA 6285  
Db |||||  
QY 5277 TCTCAGCGATCTGTCTATTTCTGTTCAATCCATAGTTGCTGACTCCCGTCTGTAGATAA 5336  
Db |||||  
QY 6286 CTACGATACGGAGGGCTTACCAATCTGCGCCAGTCTGCAATGATACCGCGAGACCCAC 6345  
Db |||||  
QY 5337 CTACGATACGGAGGGCTTACCAATCTGCGCCAGTCTGCAATGATACCGCGAGACCCAC 5396  
Db |||||  
QY 6346 GCTCACCGCTCCAGATTTATCAGCAATAAACACAGCCAGCGGAGGCGCAGAGCAGAA 6405  
Db |||||  
QY 5397 GCTCACCGCTCCAGATTTATCAGCAATAAACACAGCCAGCGGAGGCGCAGAGCAGAA 5456  
Db |||||  
QY 6406 GTGTCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTTGCGGGAAGCTAGAG 6465  
Db |||||  
QY 5457 GTGTCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTTGCGGGAAGCTAGAG 5516  
Db |||||  
QY 6466 TAAGTAGTTCCGCAAGTTAGTTGCGCAAGTTGTTGCCATTTGCTACAGGC-TCGTGG 6524  
Db |||||  
QY 5517 TAAGTAGTTCCGCAAGTTAGTTGCGCAAGTTGTTGCCATTTGCTACAGGCATCGTGG 5576  
Db |||||  
QY 6525 TGTCAAGCTCGTCTGTTGGTATGCTTCATTCAGCTCCGCTTCCCAAGATCAAGGCGAG 6584  
Db |||||  
QY 5577 TGTCAAGCTCGTCTGTTGGTATGCTTCATTCAGCTCCGCTTCCCAAGATCAAGGCGAG 5636  
Db |||||  
QY 6585 TTACATGATCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGTTG 6644  
Db |||||  
QY 5637 TTACATGATCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGTTG 5696  
Db |||||  
QY 6645 TCAGAAGTAAAGTTGGCCGAGTGTATCACTCATGTTTATGGCAGCAGCTGATTAATCTC 6704  
Db |||||  
QY 5697 TCAGAAGTAAAGTTGGCCGAGTGTATCACTCATGTTTATGGCAGCAGCTGATTAATCTC 5756  
Db |||||  
QY 6705 TTACTGTCCATGCCATCCGTTAAGATGCTTTTCTGTGACTGTGTGAGTCTCAACCAAGTCAT 6764  
Db |||||  
QY 5757 TTACTGTCCATGCCATCCGTTAAGATGCTTTTCTGTGACTGTGTGAGTCTCAACCAAGTCAT 5816  
Db |||||  
QY 6765 TCTGAGAAATAGTATGCGGAGCCGAGTTGCTCTTTCGCGCGCTCAATACGGGATAATA 6824  
Db |||||  
QY 5817 TCTGAGAAATAGTATGCGGAGCCGAGTTGCTCTTTCGCGCGCTCAACACGGGATAATA 5876  
Db |||||  
QY 6825 CCGGCCACATAGCAGAACTTTAAAGTGTCTATCATTTGAAAAAGTTCTTCGGGGCGAA 6884  
Db |||||  
QY 5877 CCGGCCACATAGCAGAACTTTAAAGTGTCTATCATTTGAAAAAGTTCTTCGGGGCGAA 5936  
Db |||||  
QY 6885 AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCCAGTTGATTAACCCACTCGTCACCCA 6944  
Db |||||  
QY 5937 AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCCAGTTGATTAACCCACTCGTCACCCA 5996  
Db |||||

QY 6945 ACTGATCTTCAGGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAAACAGGAAGGC 7004  
Db |||||  
QY 5997 ACTGATCTTCAGGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAAACAGGAAGGC 6056  
Db |||||  
QY 7005 AAAATGCCGCAAAAAAGGGAATAAGGCGCACACGGAATGTTGAATACTCTCATACTCTTCC 7064  
Db |||||  
QY 6057 AAAATGCCGCAAAAAAGGGAATAAGGCGCACACGGAATGTTGAATACTCTCATACTCTTCC 6116  
Db |||||  
QY 7065 TTTTTCATATTTAATTGAAGCATTTTATCAGGTTTATGTTCTCATGAGCGGATACATATTTG 7124  
Db |||||  
QY 6117 TTTTTCATATTTAATTGAAGCATTTTATCAGGTTTATGTTCTCATGAGCGGATACATATTTG 6176  
Db |||||  
QY 7125 AATGTATTTAGAAAAATAAACAAATAGGGTTTCCGCGCACATTTTC 7169  
Db |||||  
QY 6177 AATGTATTTAGAAAAATAAACAAATAGGGTTTCCGCGCACATTTTC 6221  
Db |||||

Search completed: May 25, 2005, 09:51:45  
Job time : 2718 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 25, 2005, 08:04:19 ; Search time 14653 Seconds  
(without alignments)  
19882.885 Million cell updates/sec

Title: US-10-006-771B-1

Perfect score: 7654

Sequence: 1 agcgtgcagctgcaggt.....gtaaaacgagcgccagtgc 7654

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_hcc:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_gss1:\*

9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1055.8	13.8	1070	1 AJ281552	AJ281552 4A3A-P6F1
2	1018.2	13.3	1048	7 CO552396	CO552396 ACly4_50
3	974.4	12.7	1013	4 BM438846	BM438846 Iplvr0015
4	947.8	12.4	1067	1 AU081137	AU081137 AU081137
5	945.8	12.4	1265	9 AG435223	AG435223 Mus. muscu
6	944.4	12.3	1027	9 CL021190	CL021190 CH216-8A1
7	924	12.1	1004	1 AJ281480	AJ281480 4A3A-P4G8
8	916.4	12.0	1089	1 AU081124	AU081124 AU081124
9	909	11.9	1338	9 AG429657	AG429657 Mus. muscu
10	893.8	11.7	1049	9 CL021189	CL021189 CH216-8A1
11	889.8	11.6	928	7 CO487414	CO487414 G00227.B7
12	882.6	11.5	889	9 CL076017	CL076017 CH216-138
13	875.6	11.4	1028	9 CL021194	CL021194 CH216-8A1
14	871.4	11.4	925	6 CB686151	CB686151 Bn01b_020
15	865	11.3	1073	7 CF269652	CF269652 Ecylocl8
16	864.4	11.3	902	7 CR753463	CR753463 DKFZp469G
17	863.8	11.3	935	4 BG838279	BG838279 Gc01_10e0
18	857.6	11.2	885	9 CL076016	CL076016 CH216-138
19	831.8	10.9	1269	9 AG393068	AG393068 Mus. muscu
20	829	10.8	841	1 AL042026	AL042026 DKFZp434E
21	820.2	10.7	865	7 CL125894	CL125894 BES182411
22	819.6	10.7	1249	8 BZ572284	BZ572284 msh2_2572
23	815.4	10.7	1025	9 CL021193	CL021193 CH216-8A1
24	815	10.6	1169	9 AG332951	AG332951 Mus. muscu

C 25	813	10.6	1284	9	AG429476	AG429476 Mus. muscu
C 26	812.6	10.6	854	4	BM438950	BM438950 Iplvr0049
C 27	807.6	10.6	966	8	BZ570738	BZ570738 msh2_1513
C 28	806.6	10.5	1321	9	AG429698	AG429698 Mus. muscu
C 29	802.4	10.5	1126	8	BZ577702	BZ577702 msh2_5533
C 30	797.4	10.4	856	7	CN823189	CN823189 Oa. splbn
C 31	796	10.4	819	6	CD649375	CD649375 CvGnd0008
C 32	785.6	10.3	1011	8	BZ576726	BZ576726 msh2_5071
C 33	785.4	10.3	1574	8	BZ572566	BZ572566 msh2_2693
C 34	783.6	10.2	820	7	CR753457	CR753457 DKFZp469P
C 35	781	10.2	827	7	CN823902	CN823902 Oa. splbn
C 36	780.8	10.2	806	4	BJ684280	BJ684280 BJ684280
C 37	779.8	10.2	846	7	CV468077	CV468077 est_1_van
C 38	779.4	10.2	1163	1	AU081044	AU081044 AU081044
C 39	778.6	10.2	833	6	CB686421	CB686421 Bn01b_04f
C 40	777.8	10.2	800	1	AJ281449	AJ281449 4A3A-P4D5
C 41	772.4	10.1	954	1	AL044364	AL044364 DKFZp434C
C 42	771.8	10.1	1245	9	AG448809	AG448809 Mus. muscu
C 43	767.6	10.0	863	7	CF752100	CF752100 TGR9 Hum
C 44	762.8	10.0	1336	8	BZ575810	BZ575810 msh2_4637
C 45	762.4	10.0	789	6	CD280920	CD280920 G44224.42

## ALIGNMENTS

RESULT 1  
AJ281552 1070 bp mRNA linear EST 30-JUN-2000  
LOCUS 4A3A-P6F11-F Anopheles gambiae immune competent 4A3A Anopheles  
DEFINITION gambiae cDNA clone 4A3A-P6F11, mRNA sequence.

ACCESSION AJ281552

VERSION AJ281552.1 GI:6929432

KEYWORDS EST.

SOURCE Anopheles gambiae (African malaria mosquito)

ORGANISM Anopheles gambiae

Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;

REFERENCE 1 (bases 1 to 1070)

AUTHORS Dimopoulos, G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C.,  
Donohue, M., Schultz, J., Bense, V., Bork, P., Ansoorge, W., Soares, M.B.  
and Kafatos, F.C.

TITLE Anopheles gambiae pilot gene discovery project: identification of  
mosquito innate immunity genes from expressed sequence tags  
generated from immune-competent cell lines

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)

MEDLINE 20300950

PUBMED 10841561

COMMENT Contact: Dimopoulos G

Fotis C. Kafatos Laboratory

European Molecular Biology Laboratory

Meyerhofstrasse 1, 69117 Heidelberg, Germany.

Location/Qualifiers

1. .1070

/organism="Anopheles gambiae"

/mol\_type="mRNA"

/strain="4A r/r"

/db\_xref="taxon:7165"

/clone="4A3A-P6F11"

/cell\_line="immune competent 4A3A"

/lab\_host="E. coli DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: EcoRI; Site\_2: NotI; sequenced from  
forward priming site which reads from the 3' end of the  
cDNA. The 4A3A is a directionally cloned and normalized  
cDNA library that was constructed from the 4A3A cell line  
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares  
(1996) : Normalization and Subtraction: Two approaches to  
Facilitate Gene Discovery, Genome Research 6, 791-806."

ORIGIN

Query Match		13.8%; Score 1055.8; DB 1; Length 1070;
Best Local Similarity		99.7%; Pred. No. 3.9e-290;
Matches 1068; Conservative		0; Mismatches 2; Indels 1; Gaps 1;
Qy	5427	TTGCTGGCGTCTTTTCCATAGGCTCGCCCCCTCGACGACATCAAAAAATCGACGCTCA 5486
Db	1	TTGCTGGCGTCTTTTCCATAGGCTCGCCCCCTCGACGACATCAAAAAATCGACGCTCA 60
Qy	5487	AGTCAGAGTGGCGAAACCCGACAGACTATTAAGATACCAAGCGTTCCTCCCTCGAAGC 5546
Db	61	AGTCAGAGTGGCGAAACCCGACAGACTATTAAGATACCAAGCGTTCCTCCCTCGAAGC 120
Qy	5547	TCCCTCGCGCTCTCTCTCCGACCTCGCGCTTACCGATACCTGTCCGCTTCTC 5606
Db	121	TCCCTCGCGCTCTCTCTCCGACCTCGCGCTTACCGATACCTGTCCGCTTCTC 180
Qy	5607	CCTTCGGAAGCGTGGCGCTTCTCATAGCTACGCTGTAGTATCTCAGTTGGTGTAG 5666
Db	181	CCTTCGGAAGCGTGGCGCTTCTCATAGCTACGCTGTAGTATCTCAGTTGGTGTAG 240
Qy	5667	GTCCGTTCGCTCCAACTGGCTGTGTCAGCAACCCCGCTTCAGCCGCGCTGCGC 5726
Db	241	GTCCGTTCGCTCCAACTGGCTGTGTCAGCAACCCCGCTTCAGCCGCGCTGCGC 300
Qy	5727	TTATCCGCTTAATCTGCTTCCGACCTCAACCGGTAAAGACAGCTTATCGCACTGGCA 5786
Db	301	TTATCCGCTTAATCTGCTTCCGACCTCAACCGGTAAAGACAGCTTATCGCACTGGCA 360
Qy	5787	CGAGCCACTGGTAAACAGGATTAGCAGACGAGGTTATGTAGCGGTGTCTACAGATTCTTG 5846
Db	361	CGAGCCACTGGTAAACAGGATTAGCAGACGAGGTTATGTAGCGGTGTCTACAGATTCTTG 420
Qy	5847	AAGTGGTGGCTTAATCTAGCGCTACACTAGAGGACAGTATTTGGTATCTGGCTCTGCTG 5906
Db	421	AAGTGGTGGCTTAATCTAGCGCTACACTAGAGGACAGTATTTGGTATCTGGCTCTGCTG 480
Qy	5907	AAGCAGATTACTCTCGGAAAAAGAGTGTGTAGCTCTTCATCGGCAACAAACACCGCT 5966
Db	481	AAGCAGATTACTCTCGGAAAAAGAGTGTGTAGCTCTTCATCGGCAACAAACACCGCT 540
Qy	5967	GGTACCGGTGTTTTTTTGTTCGAAAGCAGGATTACCGCGAGAAAAAAGGATCTCAA 6026
Db	541	GGTACCGGTGTTTTTTTGTTCGAAAGCAGGATTACCGCGAGAAAAAAGGATCTCAA 600
Qy	6027	GAAATCTTTGATCTTTTCTACGGGCTGACGCTCAGTGAAGCAAAATCAGTTAA 6086
Db	601	GAAATCTTTGATCTTTTCTACGGGCTGACGCTCAGTGAAGCAAAATCAGTTAA 660
Qy	6087	GGGATTTTGGTTCATGAGATTATCAAAAAAGGATCTTCACCTAGATCTTTTAAATTAATA 6146
Db	661	GGGATTTTGGTTCATGAGATTATCAAAAAAGGATCTTCACCTAGATCTTTTAAATTAATA 720
Qy	6147	TGAAGTTTAAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATGC 6206
Db	721	TGAAGTTTAAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATGC 780
Qy	6207	TGAATCAGTGGGACCTATCTCAGGATCTGTCTATTTTGGTTCATCATAGTTGCTGTA 6266
Db	781	TGAATCAGTGGGACCTATCTCAGGATCTGTCTATTTTGGTTCATCATAGTTGCTGTA 840
Qy	6267	CTCCCGCTGCTGATACATACGATACGAGGCGCTTACCATCTCGGCGCCAGTGTGCA 6326
Db	841	CTCCCGCTGCTGATACATACGATACGAGGCGCTTACCATCTCGGCGCCAGTGTGCA 900
Qy	6327	ATGATACCGCGAGACCCAGCTCCAGGCTCCAGATTTTATCAGCAATTAACCCAGCGCC 6386
Db	901	ATGATACCGCGAGACCCAGCTCCAGGCTCCAGATTTTATCAGCAATTAACCCAGCGCC 960
Qy	6387	GGAAGGCGGAGCGAGAGTGGTCTCGAATTTTATCCGCTTCAATCCAGTCTATTAAAT 6446
Db	961	GGAAGGCGGAGCGAGAGTGGTCTCGAATTTTATCCGCTTCAATCCAGTCTATTAAAT 1019
Qy	6447	TGTTGCCGGGAAGCTAGAGTAAGTAGTTCCGCGAGTTAATAGTTTTCGCGCAAC 6497

Db	1020	TGTTGCCGGGAAGCTAGAGTAAGTAGTTCCGCGAGTTAATAGTTTTCGCGCAAC 1070
RESULT 2		
COS52396		
LOCUS		
DEFINITION	Acly4_50 Sea lamprey Acly Petromyzon marinus cDNA, mRNA	linear EST 01-SEP-2004
ACCESSION	COS52396	
VERSION	COS52396.1	GI:51800732
KEYWORDS	EST.	
SOURCE	Petromyzon marinus (sea lamprey)	
ORGANISM	Petromyzon marinus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;	
AUTHORS	Petromyzontiformes; Petromyzontidae; Petromyzon.	
TITLE	1 (bases 1 to 1048)	
JOURNAL	Pancer, Z., Mayer, W.E., Klein, J., and Cooper, M.D.	
COMMENT	Protypic T-cell receptor and CD4-like coreceptor expressed in lymphocytes of the agnathan sea lamprey	
	Proc. Natl. Acad. Sci. U.S.A. 101, 13273-13278 (2004)	
	Contact: Pancer, Zeev	
	Division of Developmental and Clinical Immunology	
	The University of Alabama at Birmingham	
	378 Wallace Tumor Institute, 1530 Third Avenue, South, Birmingham,	
	AL 35294-3300	
	Tel: 205-975-5812	
	Fax: 205-975-7218	
	Email: zpancer@uab.edu.	
FEATURES		
source		
1. 1048		
/organism="Petromyzon marinus"		
/mol_type="mRNA"		
/db_xref="taxon:7757"		
/cell_type="lymphocyte"		
/dev_stage="immune stimulated larvae"		
/clone_lib="Sea lamprey Acly"		
/note="vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR subtracted cDNA libraries of immune stimulated larvae. All are single pass 5' or 3' sequences randomly cloned in pGEM-T Easy (Promega)."		
ORIGIN		
Query Match		
Best Local Similarity		
Matches 1043; Conservative		
1; Mismatches 2; Indels 3; Gaps 2;		
Qy	5363	TAAACGAGGAAAGAACATGTGAGCAAAAGGCGCAGCAAAAGGCGCAGGAAACCGTAAAAAGGC 60
Db	1	TAAACGAGGAAAGAACATGTGAGCAAAAGGCGCAGCAAAAGGCGCAGGAAACCGTAAAAAGGC 60
Qy	5423	CGGTTGCTGGCGTCTTTCATAGGCTCCGCGCCCGCTGACGAGCATCAAAAAATCGACG 5482
Db	61	CGGTTGCTGGCGTCTTTCATAGGCTCCG-CCGCTGACGAGCATCAAAAAATCGACG 119
Qy	5483	CTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAAGATACGAGGCTTCCCGCTGG 5542
Db	120	CTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAAGATACGAGGCTTCCCGCTGG 179
Qy	5543	AAGTCCCTCGTGGCTCTCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTCCGCTT 5602
Db	180	AAGTCCCTCGTGGCTCTCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTCCGCTT 239
Qy	5603	TCTCCCTCGGGAAGCGTGGCGCTTTCATAGCTCAGCTGAGGTATCTCAGTTCGGT 5662
Db	240	TCTCCCTCGGGAAGSGTGGCGCTTTCATAGCTCAGCTGAGGTATCTCAGTTCGGT 299
Qy	5663	GTAGTCTTCCGCTCCAGCTGGGCTGTGTGTCAGAACCCCGCTTACGCGCCGCTG 5722
Db	300	GTAGTCTTCCGCTCCAGCTGGGCTGTGTGTCAGAACCCCGCTTACGCGCCGCTG 359
Qy	5723	CGCCTTATCCGGTAATCTATCGTCTTTGAGTCAACCCCGCTTAAGACACGACTTATCGCCT 5782
Db	360	CGCCTTATCCGGTAATCTATCGTCTTTGAGTCAACCCCGCTTAAGACACGACTTATCGCCT 419





Db	174	CAGAACTTTAAAGTGTCTCATCTTGAAAAAGTTCTTCGGGGCGAAAACTCTCAAGAT	115
Qy	6998	CTTACCGCTGTGTAGATCCAGTTCGATGTAAACCACTCGTGCAACCAACTGATTTCCAGC	6957
Db	114	CTTACCGCTGTGTAGATCCAGTTCGATGTAAACCACTCGTGCAACCAACTGATTTCCAGC	55
Qy	6958	ATCTTTTAACTTCAACGACGGTTTCTGGGTGAGCAAAAACAGGAGGCAAAATGC	7011
Db	54	ATCTTTAACTTTTCAACGACGGTCTCTGGGTGAGCAAAAACAGGAGGCAAAATGC	1

RESULT 4					
AU081137					
LOCUS	AU081137	1067 bp	mRNA	linear	EST 30-JUL-2002
DEFINITION	AU081137 Oncorhynchus mykiss kidney infected by infectious hematopoietic necrosis virus Oncorhynchus mykiss cDNA clone KI2, mRNA sequence.				

VERSION	AU081137.1	GI:6431485
KEYWORDS	EST.	
SOURCE	Oncorhynchus mykiss (rainbow trout)	
ORGANISM	Oncorhynchus mykiss	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.	
REFERENCE	1. (bases 1 to 1067)	
AUTHORS	Kono T., Sakai M. and Lapetra, S.E.	
TITLE	Expressed Sequence Tag Analysis of Kidney and Gill Tissues from Rainbow Trout ( <i>Oncorhynchus mykiss</i> ) Infected with Infectious Hematopoietic Necrosis Virus	
JOURNAL	Mar. Biotechnol. 2 (5): 493-498 (2001)	

CONTACT: Masamichi Sakai  
Faculty of Agriculture  
Miyazaki University  
1-1 nishi gakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan  
Email: m.sakai@cc.miyazaki-u.ac.jp.

FEATURES

	Location/Qualifiers
SOURCE	1. 1067

```

"parentName"="mykiss", "mykissus" "mykiss"
"mol_type"="mRNA"
"db_xref"="taxon:8022"
"accession"="X01212"
"clone_k12"
"tissue type"="kidney"
"clone_11b"="Oncorhynchus mykiss Kidney infected by
infectious hematopoietic necrosis virus"
"note"="common name:rainbow trout ; infected by infectious

```

Query Match 12.4%; Score 947.8; DB 1; Length 1067;  
 Best Local Similarity 99.1%; Pred. No. 3.5e-259;  
 Matches 974; Conservative 0; Mismatches 7; Indels 2; Gaps 2;  
 QY 5026 CGTAATCATCGTCATAGCTCTTTCTCTGTGCAAAATGTTATCGCGTCACAAATTCACACA 5085  
 |||||

Qy	5326	CAAAAGCGGGTAATACCGGTTATTCACAGAATACAGGGGATTAACGACAGAAAGAAACATGTGAG	5385
Db	387	CAAAAGCGGGTAATACCGGTTATTCACAGAATACAGGGGATTAACGACAGAAAGAAACATGTGAG	446
Qy	5386	CAAAAGCGCGACGAAAGGCCAGGAAACGGTAAAGAGCCGCGTTGCTGGCGTGTTCATCA	5445
Db	447	CAAAAGCGCGACGAAAGGCCAGGAAACGGTAAAGAGCCGCGTTGCTGGCGTGTTCATCA	506
Qy	5446	GGCTCGCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGCGGAAACC	5505
Db	507	GGCTCGCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGCGGAAACC	566
Qy	5506	CGACAGGATPATAAAGATACACAGCGGTTTCCCTCGAAGCTCCCTCGTGGCTCTCCTG	5565
Db	567	CGACAGGATPATAAAGATACACAGCGGTTTCCCTCGAAGCTCCCTCGTGGCTCTCCTG	626
Qy	5566	TTCCGACCCCTGCGCTTACCGGATACGTGCGCCCTTCTCCCTCGGGAAGCGTGGCGC	5625
Db	627	TTCCGACCCCTGCGCTTACCGGATACGTGCGCCCTTCTCCCTCGGGAAGCGTGGCGC	686
Qy	5626	TTTCTCATAGCTCACGCTGTGAGTATCTCAGTTTCGGGTAGGTCGTTGGCTCCAAAGCTGG	5685
Db	687	TTTCTCATAGCTCACGCTGTGAGTATCTCAGTTTCGGGTAGGTCGTTGGCTCCAAAGCTGG	746
Qy	5686	GCTGTGTGCACGAACCCCGCTTCAGCCGACCGCTCGCTTATCCGGTAACTATCGTC	5745
Db	747	GCTGTGTGCACGAACCCCGCTTCAGCCGACCGCTCGCTTATCCGGTAACTATCGTC	806
Qy	5746	TTGAGTCCAAACCGGTGAAGACACAGCTTATCGCCACTGGCAGCAGCACCTGGTAAACAGGA	5805
Db	807	TTGAGTCCAAACCGGTGAAGACACAGCTTATCGCCACTGGCAGCAGCACCTGGTAAACAGGA	866
Qy	5806	TTAGCAGAGCGGATATGAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACG	5865
Db	867	TTAGCAGAGCGGATATGAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACG	926
Qy	5866	GCTTACACTAGAAGCAGATATTGGTATCTGCGCTCTGCTGAAGCCAGTTTACCTTCGGAA	5925
Db	927	GCTTACACTAGAAGCAGATATTGGTATCTGCGCTCTGCTGAAGCCAGTTTACCTTCGGAA	986
Qy	5926	AAAGAGTTGGTAGCTCTTGATTCGGCAAAACAAACCAACCGCTGGTAGCGGTGTTTTTTTG	5985
Db	987	AAAGAGTTGGTAGCTCTTGATTCGGCAAAACAAACCAACCGCTGGTAGCGGTGTTTTTTTG	1045
Qy	5986	TTTTCGAAGCAGCATTCACGCGC	6008

[illegible]

## COMMENT

Clones are derived from the mouse BAC library MSNg01. For BAC library availability, please contact Kuniya Abe (abe@cric.riken.jp). Tsukuba Institute, Bio Resource Center, The Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyadai, Tsukuba, 305-0074 Japan  
phone: 81-298-36-9189, fax: 81-298-36-9199  
e-mail: abe@cric.riken.jp

## PRIMERS

Sequencing : TJ

## LIBRARY

Vector : pBACe3.6

R.Site 1 : EcoRI

R.Site 2 : EcoRI

## FEATURES

## source

Location/Qualifiers  
1..1265  
/organism="Mus musculus molossinus"  
/mol\_type="genomic DNA"  
/sub\_species="molossinus"  
/db\_xref="taxon:57486"  
/clone="MSNg01-313D17.TJ"  
/sex="male"  
/tissue type="mixture of kidney and spleen"  
/clone\_lib="MSNg01 Mouse Male BAC Library"

## ORIGIN

Query Match 12.4%; Score 945.8; DB 9; Length 1265;  
Best Local Similarity 97.7%; Pred. No. 1.4e-258;  
Matches 989; Conservative 0; Mismatches 20; Indels 3; Gaps 3;

Qy 6642 TTGTGACAGTAAAGTGGCCGAGTGTATCACTCATGTTATGGCAGCACTGATTAAT 6701

Db 1096 TTTTTCAGAGTAAATGCCCGCTTTTTCACATGTTA-GGCAGCCCTGCATATC 1038

Qy 6702 CTCTTACTGTCATGCCATCCGTAAGATGCTTTCTGTGACTGGTGGTACTCAACCAAGT 6761

Db 1037 TTCTTACTGTGAGCCATCGTAGATGTTTTC-GTGACTGGGAGTACTCAACCAAGT 979

Qy 6762 CATCTCAGAAATAGTATGCGGCGCAGCGATGCTCTTGCCTGCGCGCTCAATACGGGATA 6821

Db 978 CTTTTCAGAAATAGTATGCGGCGCAGCGATGCTCTTGCCTGCGCGCTCAATACGGGATA 919

Qy 6822 ATACCGGCCA-CATACAGAACTTTAAAGTGTCTCATCTGAAAAAGTCTTCGGGG 6880

Db 918 ATACCGGCCACCATAGCAGAACTNTAAAGTGTCTCATCTGAAAAAGTCTTCGGGG 859

Qy 6881 CGAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCACTCGTGCA 6940

Db 858 CGAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCACTCGTGCA 799

Qy 6941 CCCAACTGATCTTCAGCATCTTTTACTTTCACCAAGCGTTTCTGGGTGAGCAAAACAGGA 7000

Db 798 CCCAACTGATCTTCAGCATCTTTTACTTTCACCAAGCGTTTCTGGGTGAGCAAAACAGGA 739

Qy 7001 AGGCAAAATCGGCAAAAGGGAATAGGCGCAGCAGGAAATTTGGAATCTCATCTC 7060

Db 738 AGGCAAAATCGGCAAAAGGGAATAGGCGCAGCAGGAAATTTGGAATCTCATCTC 679

Qy 7061 TTCTTTTTCATATTAATGAAGCATTTATCAGGTTTATGTTCTCATGAGCGGATACATA 7120

Db 678 TTCTTTTTCATATTAATGAAGCATTTATCAGGTTTATGTTCTCATGAGCGGATACATA 619

Qy 7121 TTTGAATGTATTTAGAAAAATAAACAATAGGGGTTCCGGCGCATATTTCCCGAAAAAGTG 7180

Db 618 TTTGAATGTATTTAGAAAAATAAACAATAGGGGTTCCGGCGCATATTTCCCGAAAAAGTG 559

Qy 7181 CCACCTGAGCTCTAAGAAACCATTTATCATGACATTAACCTATATAAAATAGCGTATC 7240

Db 558 CCACCTGAGCTCTAAGAAACCATTTATCATGACATTAACCTATATAAAATAGCGTATC 499

Qy 7241 ACAGGCCCCCTTCGCTCTCGCGCGTTTCGGGTGATGACCGGTGAACCTCTGACATGCG 7300

Db 498 ACAGGCCCCCTTCGCTCTCGCGCGTTTCGGGTGATGACCGGTGAACCTCTGACATGCG 439

## Qy

7301 CTCCCGAGAGCGTCAAGCTTGTCTGTAAAGCGATGCCGGGAGCAGCAAGCCGTCAG 7360

## Db

438 CTCCCGAGAGCGTCAAGCTTGTCTGTAAAGCGATGCCGGGAGCAGCAAGCCGTCAG 379

## Qy

7361 GGCGCGTCAAGCGGTGTTGGCGGTGTCGGGCTTAACTATGGGCATCAGAGCAG 7420

## Db

378 GGCGCGTCAAGCGGTGTTGGCGGTGTCGGGCTTAACTATGGGCATCAGAGCAG 319

## Qy

7421 ATTGTACTGAGAGTGCAACCATATGCGGTGTGAATACCGCACAGATCGTAAGGAGAAA 7480

## Db

318 ATTGTACTGAGAGTGCAACCATATGCGGTGTGAATACCGCACAGATCGTAAGGAGAAA 259

## Qy

7481 TACCGATCAGGCGCCATTCGCCATTCAGGCTGGCAACTGTTGGGAAGGCGCATCGGTG 7540

## Db

258 TACCGATCAGGCGCCATTCGCCATTCAGGCTGGCAACTGTTGGGAAGGCGCATCGGTG 199

## Qy

7541 CGGCGCTCTTCGCTATTACGCCAGCTGCGCAAGAGGGGGATGCTGCAAGGCGATTAAAT 7600

## Db

198 CGGCGCTCTTCGCTATTACGCCAGCTGCGCAAGAGGGGGATGCTGCAAGGCGATTAAAT 139

## Qy

7601 TGGGTAAAGCGCAGGGTTTCCCGAGTCAAGCTTGTAAACGACGCGCAGTG 7652

## Db

138 TGGGTAAAGCGCAGGGTTTCCCGAGTCAAGCTTGTAAACGACGCGCAGTG 87

## RESULT 6

CL021190/c

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Genome Sequencing Center

Washington University School of Medicine

Email: submissions@wustl.edu

Insert Length: 175000 Std Error: 0.00

Seq primer: Sp6 ATTTAGTGACACTATAG

Class: BAC ends

High quality sequence start: 34

High quality sequence stop: 870.

Location/Qualifiers

1..1027

/organism="Xenopus tropicalis"

/mol\_type="genomic DNA"

/strain="Nigerian frog"

/db\_xref="taxon:8364"

/clone="CH216-8A14"

/sex="male"

/cell\_line="Stock 248 F7A2, inbred N7"

/clone\_lib="CH216"

/notes="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis BAC library"

FEATURES

source

1..1027

/organism="Xenopus tropicalis"

/mol\_type="genomic DNA"

/strain="Nigerian frog"

/db\_xref="taxon:8364"

/clone="CH216-8A14"

/sex="male"

/cell\_line="Stock 248 F7A2, inbred N7"

/clone\_lib="CH216"

/notes="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis BAC library"

ORIGIN

Query Match 12.3%; Score 944.4; DB 9; Length 1027;

Best Local Similarity 98.2%; Pred. No. 3.3e-258;

Matches 976; Conservative 0; Mismatches 16; Indels 2; Gaps 2;

Qy 6646 CAGAAGTAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATATTTCT 6705

Db 1025 CAAAGTAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATATTTCT 966



```
Db 421 GGTATGAGCGGTGTCTACAGAGTTCTTGAAGTGGTGGCCCTAACTAGCGTCTACACTAGAA 480
Qy 5878 GGCAGATATTGGTATCTGCGCTCTGCTGAAGCGAGTTACCTTCGGAAAAAGAGTTGGTA 5937
Db 481 GAACAGATATTGGTATCTGCGCTCTGCTGAAGCGAGTTACCTTCGGAAAAAGAGTTGGTA 540
Qy 5938 GCTCTTGATCGGCAAAACCAACCGCTGCTGAGCGGTGTTTTTTTGGTTGCAAGCAGC 5997
Db 541 GCTCTTGATCGGCAAAACCAACCGCTGCTGAGCGGTGTTTTTTTGGTTGCAAGCAGC 600
Qy 5998 AGATTACGCGCAGAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGCGGTCTG 6057
Db 601 AGATTACGCGCAGAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGCGGTCTG 660
Qy 6058 ACCTCTCAGTGGAAACGAAACTCAGCTTAAAGGATTTTGGTCATGAGATTAATCAAAAAGGA 6117
Db 661 ACCTCTCAGTGGAAACGAAACTCAGCTTAAAGGATTTTGGTCATGAGATTAATCAAAAAGGA 720
Qy 6118 TCTTACCTAGATCTCTTTAAATTAATAAGTAAAGTTTAAATCAATCTAAAGTATATATG 6177
Db 721 TCTTACCTAGATCTCTTTAAATTAATAAGTAAAGTTTAAATCAATCTAAAGTATATATG 780
Qy 6178 AGTAAACTCTGCTGACAGTTACCAATGCTTAATCAGTGAAGGACCTATCTCAGCGATCT 6237
Db 781 AGTAAACTCTGCTGACAGTTACCAATGCTTAATCAGTGAAGGACCTATCTCAGCGATCT 839
Qy 6238 GTCTATTTCGTTTCATCATAGTTGCTGACTCCCGCTGCTGATAGATACTACGATACGGG 6297
Db 840 GTCTATTTCGTTTCATCATAGTTGCTGACTCCCGCTGCTGATAGATACTACGATACGGG 898
Qy 6298 AGGCTTACCATCTGCGCCCGAGTGTCAATGATACCGAGACCCACGCTCACCAGCTC 6357
Db 899 AGGCTTACCATCTGCGCCCGAGTGTCAATGATACCGAGACCCACGCTCACCAGCTC 955
Qy 6358 CAGATTATCAGCAATAAACCCAGCCAGCGGAGGCGCGGAGCCAGGAGTG 6408
Db 956 CAGATTATCAGCAATAAACCCAGCCAGCGGAGGCGCGGAGCCAGGAGTG 1004

RESULT 8
AU081124 1089 bp mRNA linear EST 30-JUL-2002
LOCUS AU081124 Oncorhynchus mykiss Kidney infected by infectious
DEFINITION hematopoietic necrosis virus Oncorhynchus mykiss cDNA clone KG/12,
mRNA sequence.
ACCESSION AU081124.1 GI:6431472
VERSION AU081124.1
KEYWORDS EST.
SOURCE Oncorhynchus mykiss (rainbow trout)
ORGANISM Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE 1 (bases 1 to 1089)
AUTHORS Kono,T., Sakai,M. and LaPatra,S.E.
TITLE Expressed Sequence Tag Analysis of Kidney and Gill Tissues from
Rainbow Trout (Oncorhynchus mykiss) Infected with Infectious
Hematopoietic Necrosis Virus
JOURNAL Mar. Biotechnol. 2 (5), 493-498 (2001)
COMMENT Contact: Masahiro Sakai
Faculty of Agriculture
Miyazaki University
1-1 nishi gakukenibandai, Miyazaki, Miyazaki 889-2192, Japan
Email: m.sakai@cc.miyazaki-u.ac.jp.
Location/Qualifiers
FEATURES
source
1..1089
/organism="Oncorhynchus mykiss"
/mol_type="mRNA"
/db_xref="taxon:8022"
/clone="KG/12"
/tissue_type="kidney"
/clone_lib="Oncorhynchus mykiss Kidney infected by
```

```
infectious hematopoietic necrosis virus"
/notes="common name:rainbow trout ; infected by infectious
hematopoietic necrosis virus"
ORIGIN
Query Match 12.0%; Score 916.4; DB 1; Length 1089;
Best Local Similarity 97.3%; Pred. No. 3.5e-250;
Matches 975; Conservative 0; Mismatches 21; Indels 6; Gaps 4;
Qy 5026 CGTAATCATGCTCATAGCTGTTTCTCTGTGTAATTTGTTATCGCTCACAATTTCCACACA 5085
Db 88 CGTAATCATGCTCATAGCTGTTTCTCTGTGTAATTTGTTATCGCTCACAATTTCCACACA 147
Qy 5086 ACATAGAGCGCGAAGCATAAAGCTGTAAGCTTGGGGTGCCTTAATGAGTGAGCTAACTCA 5145
Db 148 ACATAGAGCGCGAAGCATAAAGCTGTAAGCTTGGGGTGCCTTAATGAGTGAGCTAACTCA 207
Qy 5146 CATTAAATTCGTTGCGCTCACTGCCCGCTTTCAGTCGGGAAACCTGTCGTGGCAGCTGC 5205
Db 208 CATTAAATTCGTTGCGCTCACTGCCCGCTTTCAGTCGGGAAACCTGTCGTGGCAGCTGC 267
Qy 5206 ATTAATGAATTCGGCCAAACGCGCGGGAGAGCGGTTTTCGCTATTGGGCGCTCTTCCGCTT 5265
Db 268 ATTAATGAATTCGGCCAAACGCGCGGGAGAGCGGTTTTCGCTATTGGGCGCTCTTCCGCTT 327
Qy 5266 CTTGCTCTACTGACTCGCTGCGCTCGGTCGTTTCGCTGCGGCGAGCGGTATCAGCTCACT 5325
Db 328 CTTGCTCTACTGACTCGCTGCGCTCGGTCGTTTCGCTGCGGCGAGCGGTATCAGCTCACT 387
Qy 5326 CAAAGCGGTTAATACCGTTATCCACAGNATCAGGGGATACGCGGAAAGCAATGTGAG 5385
Db 388 CAAAGCGGTTAATACCGTTATCCACAGNATCAGGGGATACGCGGAAAGCAATGTGAG 447
Qy 5386 CAAAAGCCAGCAAAAGGCGAGGAAACGTAAGGCGCGGTTTTCGCTGCTTTTCCATA 5445
Db 448 CAAAAGCCAGCAAAAGGCGAGGAAACGTAAGGCGCGGTTTTCGCTGCTTTTCCATA 507
Qy 5446 GGCTCCGCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAAGAGTGGCGAAACC 5505
Db 508 GGCTCCGCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAAGAGTGGCGAAACC 567
Qy 5506 CGACAGACTATAAAGATACCGAGCTTTCCCGCTGGAAGCTCCCTGTCGCTCTCTCTG 5565
Db 568 CGACAGACTATAAAGATACCGAGCTTTCCCGCTGGAAGCTCCCTGTCGCTCTCTCTG 627
Qy 5566 TTCCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTGGCGC 5625
Db 628 TTCCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTGGCGC 687
Qy 5626 TTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGTTGAGGTGCTTCGCTCCAAGCTGG 5685
Db 688 TTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGTTGAGGTGCTTCGCTCCAAGCTGG 747
Qy 5686 GCTGTGTGACGAACCCCGCTTACCGCGACCGCTGCGCTTATCCGGTAACCTATCGTC 5745
Db 748 GCTGTGTGACGAACCCCGCTTACCGCGACCGCTGCGCTTATCCGGTAACCTATCGTC 807
Qy 5746 TTGAGTCCAAACCCCGGTAAGACACGACTATTCGCACTGGCAGCAGCCACCTGTTAAACAGGA 5805
Db 808 TTGAGTCCAAACCCCGGTAAGACACGACTATTCGCACTGGCAGCAGCCACCTGTTAAACAGGA 867
Qy 5806 TTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTTGAAG-TGGTGGCCCTAACTAC 5864
Db 868 TTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTTGAAGTTTGTGGCTTAACCTAC 927
Qy 5865 GGCTACATAGAAAGGACAGTATTTGGTATCTGCGCTCTGCTGAGCCAGGTACCTTCGGA 5924
Db 928 GGCTACATAGAAAGGACAGTATTTGGTATCTGCGCTCTGCTGAGCCAGGTACCTTCGGA 987
Qy 5925 AAAAGAGTTGGTAGCTCTT-GATCCCGCAACCAACCA---CCGCTGGTAGCGGTGTTT 5980
Db 988 AAAAGAGTTGGTAGCTCTTGGATCCCGCAACCAACCAACCCGCTGGGAAGGGGGTGTCT 1047
```



```

Qy 5981 TTTTGTGTCAGCA-GCAGATTACGCGCAGAAAAAAGGAT 6021
      |||||
Db 1048 TTTTGTGTCAGCAGCAGATTACCGGAAAAACGGAT 1089

RESULT 9
AG429657/c 1338 bp DNA linear GSS 03-JUN-2004
DEFINITION Mus musculus molossinus DNA, clone:MSMg01-304F12.TJ, genomic survey
sequence.
ACCESSION AG429657
VERSION AG429657.1 GI:48072720
KEYWORDS GSS.
SOURCE Mus musculus molossinus
ORGANISM Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
BAC end Sequences of Library MSMg01
Unpublished
2 (bases 1 to 1338)
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
Direct Submission
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Japan
1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center.
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY : pBACe3.6
Vector : EcoRI.
R.Site 1 : EcoRI.
R.Site 2 : EcoRI.

FEATURES
    source
    1..1338 Location/Qualifiers
        /organism="Mus musculus molossinus"
        /mol_type="genomic DNA"
        /sub_species="molossinus"
        /db_xref="taxon:57486"
        /clone="MSMg01-304F12.TJ"
        /sex="male"
        /tissue_type="mixture of kidney and spleen"
        /clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN
    Query Match 11.9%; Score 909; DB 9; Length 1338;
    Best Local Similarity 92.1%; Pred.No.5.1e-248;
    Matches 1025; Conservative 0; Mismatches 77; Indels 11; Gaps 6;

Qy 6540 TTGGTATGGCTTCATTCAGCTCCGGTTCCTCCCAACGATCAAGGCGAGTTACATGATCCCCCA 6599
      |||||
Db 1191 TTTTGTGTTGAAGGTTCTCCTACTCGGTTCCCAAGGATCAAGGGGTTCCCTGTTCCTCCCC- 1133

Qy 6600 TGTGTGCAAAAACGGTTAGTCTCCTTCGGTCTCCGATCGTTGTGTCAGAGTAAGTTGG 6659
      |||||
Db 1132 ---TTTGACAAAAAGGGGTTCTTCTTCGTCTCCNATTC--TTTCCAAAAGTTAAATTGC 1078

Qy 6660 CGCAGTGTATCACTCATGTTATGCGAGCACTGCATAATTCTCTTACTGTGTCATGCCAT 6719
      |||||
Db 1077 CGCA--GTTTTCTTCAGGTAGGCGAGCCCGCTTATTTCTTACTGCCACCCAT 1020

Qy 6720 CGGTAAGATGC-TTTTCTGTGATCGGTGAGTACTCAA-CCAAGTCATTCTGAGAAATAGTG 6777
      |||||
Db 1019 CCATAAAGATGCTTTTCTGTGCGGGGGGAGTACTCAACCAAGTCTTCTCTGAGAATAGTG 960

```

```

Qy 6778 TATCGGCGACCGAGTTGCTCTTGC CGGCGTCAATACGGGATAATACCGCGCCACATAG 6837
      |||||
Db 959 TATCGGCGACCGAGTTGCTCTTGC CGGCGTCAATACGGGATAATACCGCGCCACATAG 900
      |||||
Qy 6838 CAGAACTTTTAAAGTGTCTCATTTGGAACGTTTCTTCGGGGCGAAACCTCTCAAGAT 6897
      |||||
Db 899 CAGAACTNTAAAGTGTCTCATTTGGAACGTTTCTTCGGGGCGAAACCTCTCAAGAT 840
      |||||
Qy 6998 CTTACCGCTGTTGAGATCCAGTTGATGTAACCACTGCTGCACCAACTGATCTTTCAGC 6957
      |||||
Db 839 CTTACCGCTGTTGAGATCCAGTTGATGTAACCACTGCTGCACCAACTGATCTTTCAGC 780
      |||||
Qy 6958 ATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAACAGAAAGGCAAAATCGCGCAA 7017
      |||||
Db 779 ATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAACAGAAAGGCAAAATCGCGCAA 720
      |||||
Qy 7018 AAAGGGAATAGGGCGACACGGAATGTTGAATACTCATACTCTTCTCTTTTCAATATTA 7077
      |||||
Db 719 AAAGGGAATAGGGCGACACGGAATGTTGAATACTCATACTCTTCTCTTTTCAATATTA 660
      |||||
Qy 7078 TTGAAGCATTTATCAGGGTTATTGTTCTCATGAGCGGATACATATTGTAATGTATTAGAA 7137
      |||||
Db 659 TTGAAGCATTTATCAGGGTTATTGTTCTCATGAGCGGATACATATTGTAATGTATTAGAA 600
      |||||
Qy 7138 AAATAAACAATAGGGTTCGCGCACATTTCCCGCAAAAGTGCCACCTGACGCTCTAGA 7197
      |||||
Db 599 AAATAAACAATAGGGTTCGCGCACATTTCCCGCAAAAGTGCCACCTGACGCTCTAGA 540
      |||||
Qy 7198 AACCATTTATTCATGACATTAACCTATAAAATAGGGTATACAGAGCCCTTTCTGCT 7257
      |||||
Db 539 AACCATTTATTCATGACATTAACCTATAAAATAGGGTATACAGAGCCCTTTCTGCT 480
      |||||
Qy 7258 CGCGCGTTTCGGTGATGACCGTGAAAACCTCTGACATGACGCTCCCGGAGACGGTCA 7317
      |||||
Db 479 CGCGCGTTTCGGTGATGACCGTGAAAACCTCTGACATGACGCTCCCGGAGACGGTCA 420
      |||||
Qy 7318 AGCTTGTCTGTAAGCGGATGCGGAGCAGACAAAGCCGTCAGGCGCGTCCAGCGGGT 7377
      |||||
Db 419 AGCTTGTCTGTAAGCGGATGCGGAGCAGACAAAGCCGTCAGGCGCGTCCAGCGGGT 360
      |||||
Qy 7378 TGGCGGGTGTCCGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGA 7437
      |||||
Db 359 TGGCGGGTGTCCGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGA 300
      |||||
Qy 7438 CCATATGCGGTGTGAAATACCGCACAGATGCTTAAGGAGAAATAACGCATCAGCGCCA 7497
      |||||
Db 299 CCATATGCGGTGTGAAATACCGCACAGATGCTTAAGGAGAAATAACGCATCAGCGCCA 240
      |||||
Qy 7498 TTGCGCATTCAGGCTGCGCACTGTTGGGAGGGGATCGGTGCGGGCTCTTCGCTATT 7557
      |||||
Db 239 TTGCGCATTCAGGCTGCGCACTGTTGGGAGGGGATCGGTGCGGGCTCTTCGCTATT 180
      |||||
Qy 7558 ACGCAGCTGCGGAAAAGGGGATGCTGCTGAAGCGGATTAAGTTGGGTAAACGCCA-CGGT 7616
      |||||
Db 179 ACGCAGCTGCGGAAAAGGGGATGCTGCTGAAGCGGATTAAGTTGGGTAAACGCCAGGGT 120
      |||||
Qy 7617 TTTCCCATGACGAGTTGTAAACACGCGCCA 7649
      |||||
Db 119 TTTCCCATGACGAGTTGTAAACACGCGCCA 87
      |||||

```

```

RESULT 10
CL021189 LOCUS 1049 bp DNA linear GSS 31-DEC-2003
DEFINITION CH216-8A14 RM1.1 CH216 xenopus tropicalis genomic clone CH216-8A14,
genomic survey sequence.
ACCESSION CL021189
VERSION CL021189.1 GI:40463002
KEYWORDS GSS.
SOURCE xenopus tropicalis (western clawed frog)
ORGANISM xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus; Silurana.	
1 (bases 1 to 1049)	
Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T., Mardis,E. and Wilson,R.	
A physical map of the xenopus tropicalis genome	
Unpublished (2003)	
Contact: Richard K Wilson	
Genome Sequencing Center	
Washington University School of Medicine	
Email: submissions@watson.wustl.edu	
Insert Length: 175000 Std Error: 0.00	
Seq primer: RM1 TAGACTCATCTATAGGGAGA	
Class: BAC ends	
High quality sequence start: 43	
High quality sequence stop: 888.	
Location/Qualifiers	
1..1049	
/organism="Xenopus tropicalis"	
/mol_type="genomic DNA"	
/strain="Nigerian frog"	
/db_xref="taxon:8364"	
/clone="CH216-8A14"	
/sex="male"	
/cell_line="Stock 248 F7A2, inbred N7"	
/clone_lib="CH216"	
/note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis BAC library"	
ORIGIN	
Query Match 11.7%; Score 893.8; DB 9; Length 1049;	
Best Local Similarity 98.5%; Pred. No. 1.le-243;	
Matches 923; Conservative 0; Mismatches 12; Indels 2; Gaps 2;	
5026 CGTAATCATGTCATAGCTGTTCCCTGTGTGAATTTGTTATCCGCTCACAAATTCACACA 5085	
112 CGTAATCATGTCATAGCTGTTCCCTGTGTGAATTTGTTATCCGCTCACAAATTCACACA 171	
5086 ACATACGAGCCGGGAAGCATAAAGTGTAAAGCCTCGGGTGCCTAATGAGTGAGCTAACTCA 5145	
172 ACATACGAGCCGGGAAGCATAAAGTGTAAAGCTGGGTGCTTANTGAGTGAGCTAACTCA 231	
5146 CATTAAATTGGTTGGCTCACTGCCCGGTTTCAGTCGGGAAACCTGTGTGCCAGCTGC 5205	
232 CATTAAATTGGTTGGCTCACTGCCCGCTTCACAGTCGGGAAACCTGTGTGCCAGCTGC 291	
5206 ATTAATGAATCGGCCCAAGCGCGGGGAGAGCGGTTTTCGTTATTTGGGGGCTCTCCGCTT 5265	
292 ATTAATGAATCGGCCCAAGCGCGGGGAGAGCGGTTTTCGTTATTTGGGGGCTCTCCGCTT 351	
5266 CCTCGCTCACTGACTCGCTGGGCTCGGTCGTTTCGGCTCGCGCAGCGGTATCAGCTCACT 5325	
352 CCTCGCTCACTGACTCGCTGGGCTCGGTCGTTTCGGCTCGCGCAGCGGTATCAGCTCACT 411	
5326 CAAAGGCGGTAATACGGTTATTCACAGAATCAGGGGATAACGCGAGGAAAGAAACATGTGAG 5385	
412 CAAAGGCGGTAATACGGTTATTCACAGAAATCAGGGGATAACGCGAGGAAAGAAACATGTGAG 471	
5386 CAAAGGCCAGCAAAAGCCAGGAAACCGTAAAGAGCCGGTGTCTCGCGTTTTTTCATA 5445	
472 CAAAGGCCAGCAAAAGGCCAGGAAACCGTAAAGAGCCGGTGTCTCGCGTTTTTTCATA 531	
5446 GGCTCCGCCCCCTCAGCAGCATCACAATAATCGACGCTCAAGTCACAGGTGGCGAAACC 5505	
532 GGCTCCGCCCCCTCAGCAGCATCACAATAATCGACGCTCAAGTCACAGGTGGCGAAACC 591	
5506 CGACAGGACTATAAAGATAACAGCGGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCCTG 5565	
592 CGACAGGACTATAAAGATAACAGCGGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCCTG 651	
5566 TTCCGACCGCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCCGGGAAGCGTGGCGC 5625	
652 TTCCGACCGCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCCGGGAAGCGTGGCGC 711	

/note="Organ: Roots from 9 year old trees measuring approximately 4.5 m tall, and 10 cm in diameter.; Vector: pBluescript II SK (+) XR; Site 1: Eco-RI; Site 2: Xho-I; cDNA was prepared from 5 mg of poly A+ selected RNA and was directionally ligated into the pBluescript II SK (+) XR vector (Stratagene), transformed by electroporation into DH10B cells (in vitro) for propagation"

```
ORIGIN
Query Match      11.6%; Score 889.8; DB 7; Length 928;
Best Local Similarity 96.0%; Pred. No. 1.4e-242;
Matches 891; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 5438 TTTCATAGGCTCCGCCCTTGACGAGCATCACAAAATCGACGCTCAAGCTCAGAGTG 5497
Db 1 TTTTCATAGGCTCCGCCCTTGACGAGCATCACAAAATCGACGCTCAAGTNNNNNG 60

Qy 5498 GCGAAACCGACAGGACTATAAGATACAGGCGTTTCCCTGGAAGCTCCCTCGTCG 5557
Db 61 GCGAAACCGACAGGACTATAAGATACAGGCGTTTCCCTGGAAGCTCCCTCGTCG 120

Qy 5558 CTCTCCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAG 5617
Db 121 CTCTCCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAG 180

Qy 5618 CGTGGCGTTTCTATAGCTACGCTGTAGGTATCTCAGTTTCGGGTAGGTGCTTCGCTC 5677
Db 181 CGTGGCGTTTCTATAGCTACGCTGTAGGTATCTCAGTTTCGGGTAGGTGCTTCGCTC 240

Qy 5678 CAAGCTGGGCTGTGTGACGAAACCCCGTTTCAGCCGACGCTGCGCTTATCGGTAA 5737
Db 241 CAAGCTGGGCTGTGTGACGAAACCCCGTTTCAGCCGACGCTGCGCTTATCGGTAA 300

Qy 5738 CTATCGTCTTGAGTCAACCCGGTAAGACACGACTATTCGCACTGGCGAGCAGCACTGG 5797
Db 301 CTATCGTCTTGAGTCAACCCGGTAAGACACGACTATTCGCACTGGCGAGCAGCACTGG 360

Qy 5798 TAAACAGGATTAGCAGGCGAGGTATGTAGGCGGTCTACAGAGTCTTGAAGTGGTGCC 5857
Db 361 TAAACAGGATTAGCAGGCGAGGTATGTAGGCGGTCTACAGAGTCTTGAAGTGGTGCC 420

Qy 5858 TAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTCGCTCTGCTGAAAGCCAGTTAC 5917
Db 421 TAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTCGCTCTGCTGAAAGCCAGTTAC 480

Qy 5918 CTTGGGAAAAGAGTTGTAGTCTTTGATTCGGGCAACAAACCCAGCTGGTAGCGGTGG 5977
Db 481 CTTGGGAAAAGAGTTGTAGTCTTTGATTCGGGCAACAAACCCAGCTGGTAGCGGTGG 540

Qy 5978 TTTTTCCTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGATCTCAAGAAAGATCCTTT 6037
Db 541 TTTTTCCTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGATCTCAAGAAAGATCCTTT 600

Qy 6038 GATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAACCTCACGTTAAGGGATTTGGT 6097
Db 601 GATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAACCTCACGTTAAGGGATTTGGT 660

Qy 6098 CATGAGATTACAAAAGAGATCTTACCTAGATCTCTTTTAAATTAAGATGAAGTTTAA 6157
Db 661 CATGAGATTACAAAAGAGATCTTACCTAGATCTCTTTTAAATTAAGATGAAGTTTAA 720

Qy 6158 ATCAATCTAAAGTATATATAGTAAATTTGGTCTGACAGATTACCAATGCTTAATCAGTGA 6217
Db 721 ATCAATCTAAAGTATATATAGTAAATTTGGTCTGACAGATTACCAATGCTTAATCAGTGA 780

Qy 6218 GGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCTGACTCCCGCTCGT 6277
Db 781 GGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCTGACTCCCGCTCGT 840

Qy 6278 GTAGATACTACGATACGGGAGGCTTACCATCTGCGCCCGAGTGTGCAATGATACCGCG 6337
Db 841 GTNNNTACTACGNTACGGGAGGCTTACCANCTGNNNNNGTGTGCAATGATACNCGG 900
```

```
Qy 6338 AGACCCACGCTCACCAGCTCCAGATTTA 6365
Db 901 AGACCCACGCTCACCAGCTCCAGATTTA 928

RESULT 12
LOCUS CL076017/c
DEFINITION CH216-138F20_Sp6.1 CH216 Xenopus tropicalis genomic clone
ACCESSION CH216-138F20_genomic survey sequence.
VERSION CL076017
KEYWORDS CL076017.1 GI:40531930
SOURCE GSS.
ORGANISM Xenopus tropicalis (western clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 889)
AUTHORS Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
Mardis,E. and Wilson,R.
TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 175000 Std Error: 0.00
Seq primer: Sp6 ATTTAGTGACACTATAG
Class: BAC ends
High quality sequence start: 2
High quality sequence stop: 847.
FEATURES
Location/Qualifiers
1..889
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-138F20"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
BAC library"

ORIGIN
Query Match 11.5%; Score 882.6; DB 9; Length 889;
Best Local Similarity 99.6%; Pred. No. 1.6e-240;
Matches 885; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 6645 TCAGAGTAAAGTTGCGCGCAGTGTTATCACTCATGTTATGCGAGCAGCTGATATCTC 6704
Db 889 TCAAAGTAGTTTGGCGCAGTGTTATCACTCATGTTATGCGAGCAGCTGATATCTC 830

Qy 6705 TTACTGTCATGCCATCCGTAAGATGCTTTCTGTGACTGGTGAGTACTCAACCAAGTCAT 6764
Db 829 TTACTGTCATGCCATCCGTAAGATGCTTTCTGTGACTGGTGAGTACTCAACCAAGTCAT 770

Qy 6765 TCTGAGAAATAGTATGCGGCGACCGAGTTGCTCTTGCCCGCGCTCAATACGGATAATA 6824
Db 769 TCTGAGAAATAGTATGCGGCGACCGAGTTGCTCTTGCCCGCGCTCAATACGGATAATA 710

Qy 6825 CCGCGCCACATAGCAGAACTTTAAAGTCTCATCTTTGAAAGAGTTCTTTCGGGCGCAA 6884
Db 709 CCGCGCCACATAGCAGAACTTTAAAGTCTCATCTTTGAAAGAGTTCTTTCGGGCGCAA 650

Qy 6885 AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATTAACCCACTCGTGCACCCA 6944
Db 649 AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATTAACCCACTCGTGCACCCA 590

Qy 6945 ACTGATCTTACGATCTTTTACTCTTCCACGCGTTTCTGGGTGAGCAAAAAACAGGAAGC 7004
Db 589 ACTGATCTTACGATCTTTTACTCTTCCACGCGTTTCTGGGTGAGCAAAAAACAGGAAGC 530
```

```
Qy 7005 AAAATGCCGCAAAAAGGGAATAAGGCGACACGGAATGTTGAATACTCATACTCTTCC 7064
Db 529 AAAATGCCGCAAAAAGGGAATAAGGCGACACGGAATGTTGAATACTCATACTCTTCC 470
Qy 7065 TTTTTCATATATTGTAAGCAATTTATCAGGGTATTCTCTCATGAGCGGATACATATTG 7124
Db 469 TTTTTCATATATTGTAAGCAATTTATCAGGGTATTCTCTCATGAGCGGATACATATTG 410
Qy 7125 AATGATTTTGAAGAAATAAACAATAGGGTTCGCGCAATTTCCCGAAAGTGCCAC 7184
Db 409 AATGATTTTGAAGAAATAAACAATAGGGTTCGCGCAATTTCCCGAAAGTGCCAC 350
Qy 7185 CTGACGCTTAAGAAACCATATTATCATGACATTAACCTATAAATAAGCGGTATCAGA 7244
Db 349 CTGACGCTTAAGAAACCATATTATCATGACATTAACCTATAAATAAGCGGTATCAGA 290
Qy 7245 GGCCCTTTCTGTCGCGGTTTCGGTGATGACGGTGAACCTCTGACATGCAAGCTCC 7304
Db 289 GGCCCTTTCTGTCGCGGTTTCGGTGATGACGGTGAACCTCTGACATGCAAGCTCC 230
Qy 7305 CGGAGACGGTCAACAGCTTGTCTGAAGCGGATGCGGAGCAGACAAGCCGTCAGGGCG 7364
Db 229 CGGAGACGGTCAACAGCTTGTCTGAAGCGGATGCGGAGCAGACAAGCCGTCAGGGCG 170
Qy 7365 CGTCAGCGGGTGTGGCGGGTGTGCGGGCTGGCTTAACCTATGCGGCATCAGAGCAGATTG 7424
Db 169 CGTCAGCGGGTGTGGCGGGTGTGCGGGCTGGCTTAACCTATGCGGCATCAGAGCAGATTG 110
Qy 7425 TACTGAGAGTGACCATATGCGGTGTGAATACCGGACAGATGCGGTGAAGGAGAAATACC 7484
Db 109 TACTGAGAGTGACCATATGCGGTGTGAATACCGGACAGATGCGGTGAAGGAGAAATACC 50
Qy 7485 GCATCAGCGGCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCG 7533
Db 49 GCATCAGCGGCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCG 1
```

## RESULT 13

LOCUS CL021194/c 1028 bp DNA linear GSS 31-DEC-2003  
DEFINITION CH216-8A16\_Sp6.1 CH216 Xenopus tropicalis genomic clone CH216-8A16,  
genomic survey sequence.

ACCESSION CL021194

VERSION CL021194.1

KEYWORDS GI:40463007

SOURCE GSS.

ORGANISM Xenopus tropicalis (western clawed frog)

Xenopus tropicalis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;

Xenopodinae; Xenopus; Silurana.

1 (bases 1 to 1028)

Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,

Mardis, E., and Wilson, R.

A physical map of the xenopus tropicalis genome

Unpublished (2003)

Contact: Richard K Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: submissions@wustl.edu

Insert Length: 175000 Std Error: 0.00

Seq primer: Sp6 ATTTAGGTGACACTATAG

Class: BAC ends

High quality sequence start: 64

High quality sequence stop: 844.

Location/Qualifiers

1..1028

/organism="Xenopus tropicalis"

/mol\_type="genomic DNA"

/strain="Nigerian frog"

/db\_xref="taxon:9364"

/clone="CH216-8A16"

/sex="male"

## FEATURES

source

ORIGIN

Query Match 11.4%; Score 875.6; DB 9; Length 1028;  
Best Local Similarity 96.1%; Pred. No. 1.7e-238;  
Matches 930; Conservative 0; Mismatches 34; Indels 4; Gaps 3;

/cell\_line="Stock 248 F7A2, inbred N7"  
/clone\_lib="CH216"  
/notes="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis  
BAC library"

Qy 6684 TGCAGCAGTCATATAATCTCTTACTGTCTATGTCATGCGTAAAGATGCTTTTCTGTGACTG 6743  
Db 978 TGCAGCAAAATCTTCTTACTGTTTCATGCGCATCCGTAAGATGCT-TTTTTCGAGAACTG 920  
Qy 6744 GTGAGTACTCAACCAAGTCATTTCTGAGAAATAGTGTATGCGGCGACCGAGTGTCTTTGCC 6803  
Db 919 GTGAGTACTCAACCAAGTCATTTCTGAGAAATAGTGTATGCGGCGACCGAGTGTCTTTGCC 860  
Qy 6804 CGCGGTCAATACCGGATATATACCGGCGCACATAGCAGAACTTTAAAGTGTCTCATATTG 6863  
Db 859 CGCGGTCAATACCGGATATATACCGGCGCACATAGCAGAACTTTAAAGTGTCTCATATTG 800  
Qy 6864 GAAAGCTTCTTCGCGGCGGAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTTCA 6923  
Db 799 GATAAGCTTCTTCGCGGCGGAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTTCA 740  
Qy 6924 TGTAAACCCACTCGTGACACCAACTGATCTTCAGCATCTTTTACTTTTCAACAGCGTTTCTG 6983  
Db 739 TGTAAACCCACTCGTGACACCAACTGATCTTCAGCATCTTTTACTTTTCAACAGCGTTTCTG 680  
Qy 6984 GTTGAGCAAAAAACGAGAAAGCAAAATGCGCAAAAAAGGGAATAAGGCGCACACGGAAT 7043  
Db 679 GTTGAGCAAAAAACGAGAAAGCAAAATGCGCAAAAAAGGGAATAAGGCGCACACGGAAT 620  
Qy 7044 GTTGAAATCTCATACTCTTCTCTTTTCAATATTTATGAAGCATTTATCAGGGTTATTGTC 7103  
Db 619 GTTGAAATCTCATACTCTTCTCTTTTCAATATTTATGAAGCATTTATCAGGGTTATTGTC 560  
Qy 7104 TCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAAATAAGAGGGTTCGCGCA 7163  
Db 559 TCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAAATAAGAGGGTTCGCGCA 500  
Qy 7164 CATTTCCCGAAAAAGTGCCACCTGACGCTCTAAGAAACCATTTATCATGACATTAACCT 7223  
Db 499 CATTTCCCGAAAAAGTGCCACCTGACGCTCTAAGAAACCATTTATCATGACATTAACCT 440  
Qy 7224 ATAAAAATAGCGGTATCAGAGCGCTTTCGTCTCGCGCGCTTTCGTGATGACGGTCAAA 7283  
Db 439 ATAAAAATAGCGGTATCAGAGCGCTTTCGTCTCGCGCGCTTTCGTGATGACGGTCAAA 380  
Qy 7284 ACCTCTGACACATGACGCTCCCGAGACGCTCACAGCTTCTGTAGAGCGGATGCCGGGA 7343  
Db 379 ACCTCTGACACATGACGCTCCCGAGACGCTCACAGCTTCTGTAGAGCGGATGCCGGGA 320  
Qy 7344 GCAGACAAGCCGTCAGGGCGCTGACGCGGTGTGCGGGGTGTGCGGGTGTGGCTTAACT 7403  
Db 319 GCAGACAAGCCGTCAGGGCGCTGACGCGGTGTGCGGGGTGTGCGGGTGTGGCTTAACT 260  
Qy 7404 ATCGCGCATCAGACGAGATTTACTGAGAGTGACACATATGCGGTGTGAAATACCGCACA 7463  
Db 259 ATCGCGCATCAGACGAGATTTACTGAGAGTGACACATATGCGGTGTGAAATACCGCACA 200  
Qy 7464 GATGCGTAAAGGAGAAATACCGCATCAGGGCGCATTCGCCATTCAGGCTCGCAACTGTT 7523  
Db 199 GATGCGTAAAGGAGAAATACCGCATCAGGGCGCATTCGCCATTCAGGCTCGCAACTGTT 140  
Qy 7524 GGAAGGGCGCATCGGTGCGGGCTCTTCGTATTTAGCCAGCTGGCGAAAGGGGATGTG 7583  
Db 139 GGAAGGGCGATCGGTGCGGGCTCTTCGTATTTAGCCAGCTGGCGAAATGGGATGTG 80  
Qy 7584 CTGCAAGGGCATTAAGTGTGGTAAACCGCAGGGTTTTCACAGTGTGTAACAGTGTGTAACAG 7643  
Db 79 CTGCAAGGGCATTAAGAT-GGTAAACCAAGGGTTCCTCCAGTCAAGACGT--GAAACGA 23

```

QY      7644  CGGCCAGT 7651
Db      22  CGGCCATT 15

RESULT 14
CB686151/c
LOCUS
DEFINITION
    Bn01b_AAPC_ECORC_transgenic_Brassica_napus_overexpressing_BNCF17_c
    onstitutively_frost_tolerant_Brassica_napus_cDNA_clone_Bn01b_02o08,
    mRNA sequence.
ACCESSION
CB686151
VERSION
CB686151.1
KEYWORDS
EST.
SOURCE
    Brassica napus (rape)
ORGANISM
    Brassica napus
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
    Singh, J., Allard, G., Tinker, N., Robert, L., Lacroix, C., De Moors, A.,
    Chagnon, J., Farah, S., Couroux, P. and Hattori, J.
    Expressed Sequence Tags from Constitutively frost tolerant
    transgenic Brassica napus overexpressing BNCF17
JOURNAL
    Unpublished (2002)
COMMENT
    Contact: Singh, J.A.
    Eastern Cereal and Oilseed Research Centre
    Agriculture and Agri-food Canada
    KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, KIA
    0C6, Canada
    Tel: (613) 759-1662
    Fax: (613) 759-1701
    Email: singhja@agr.gc.ca.
FEATURES
    source
        1..925
            /organism="Brassica napus"
            /mol_type="mRNA"
            /cultivar="Westar"
            /db_xref="taxon:3708"
            /clone="Bn01b_02o08"
            /tissue_type="fourth leaf"
            /dev_stage="3 weeks seedling grown at room temperature"
            /clone_lib="Bn01b_AAPC_ECORC_transgenic_Brassica_napus_ove
            rexpressing_BNCF17_constitutively_frost_tolerant"
            /note="Vector: Bluescript SK+/XhoI-EcoRI; Site 1: EcoRI;
            Site 2: XhoI; Germinated in soil flats and seedlings grown
            for 3 weeks in a Conviron E-15 cabinet set at 20oC /16 hr
            light (250 Em-2sec-1) and 16 oC / 8 hr dark. Fourth leaves
            collected at 9 am and immediately frozen."
ORIGIN
    Query Match      11.4%; Score 871.4; DB 6; Length 925;
    Best Local Similarity 99.0%; Pred. No. 2.6e-237;
    Matches 880; Conservative 6; Mismatches 2; Indels 1; Gaps 1;

QY      5026  CCGTATCATGTCATAGCTGTTTCTGTCGTAAGTGTATCCGCTCAAAATTCACACA 5085
Db      888  CCGTATCATGTCATAGCTGTTT-CTGTGTGAATTTGTTATCCGCTCAAAATTCACACA 830

QY      5086  ACATACGAGCCGAGACATAAGTCTAAAGCTCGGGTGCCCTAATGAGTGAAGTAACTCA 5145
Db      829  ACATACGAGCCGAGACATAAGTCTAAAGCTCGGGTGCCCTAATGAGTGAAGTAACTCA 770

QY      5146  CATTAAATGCGTTGGCTCACTGCCCGCTTTTCAGTCGGGAACCTGTCGTCACAGTCGC 5205
Db      769  CATTAAATGCGTTGGCTCACTGCCCGCTTTTCAGTCGGGAACCTGTCGTCACAGTCGC 710

QY      5206  ATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTGGTATTGGGCGCTCTCCGCTT 5265
Db      709  ATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTGGTATTGGGCGCTCTCCGCTT 650

```

```

5266  CCTCGCTCACTGACTCGCTCGCTCGGTTCGGTTCGGCTCGCGGAGCGGTATCAGCTCACT 5325
      |||||
649   CCTCGCTCACTGACTCGCTCGCTCGGTTCGGTTCGGCTCGCGGAGCGGTATCAGCTCACT 590
      |||||
5326  CAAAGGCGGTAAATACGGTTATCCACAGAATCAGGGGATAACGCGAGGAAGAACATGTGAG 5385
      |||||
589   CAAAGGCGGTAAATACGGTTATCCACAGAATCAGGGGATAACGCGAGGAAGAACATGTGAG 530
      |||||
5386  CAAAGGCCACAAAAGGCCAGGACCGTAAAGGCCGGTTCGTCGCGCTTTTCCATA 5445
      |||||
529   CAAAGGCCACAAAAGGCCAGGACCGTAAAGGCCGGTTCGTCGCGCTTTTCCATA 470
      |||||
5446  GGCTCCGCCCTCCCTGACGAGCATCACAAAATCAAGCTCAAGTCAGAGGTGGCGAAACC 5505
      |||||
469   GGCTCCGCCCTCCCTGACGAGCATCACAAAATCAAGCTCAAGTCAGAGGTGGCGAAACC 410
      |||||
5506  CGACAGGACTATAAAGATACCAAGCGTTTCCCTTGGAAAGCTCCCTCTCGGCTCTCTG 5565
      |||||
409   CGACAGGACTATAAAGATACCAAGCGTTTCCCTTGGAAAGCTCCCTCTCGGCTCTCTG 350
      |||||
5566  TTCCGACCTCGCGCTTACCGGATACCTGTCGGCTTCTCCCTCGGGAAGCGTGGCGC 5625
      |||||
349   TTCCGACCTCGCGCTTACCGGATACCTGTCGGCTTCTCCCTCGGGAAGCGTGGCGC 290
      |||||
5626  TTTCTCATAGCTCAGCGTGTAGGTATCTCAGTTCGGTGTAGGTCTCGCTCCAAAGCTGG 5685
      |||||
289   TTTCTCATAGCTCAGCGTGTAGGTATCTCAGTTCGGTGTAGGTCTCGCTCCAAAGCTGG 230
      |||||
5686  GCTGTGTGCACGAACCCCGCTTTCAGCGCCAGCCCTGCGCTTATTCGGTAACTATCGTC 5745
      |||||
229   GCTGTGTGCACGAACCCCGCTTTCAGCGCCAGCCCTGCGCTTATTCGGTAACTATCGTC 170
      |||||
5746  TTGAGTCCAAACCCCGTAAAGACACCACTTATCGCACCTGGCAGCAGCCACTGGTAACAGA 5805
      |||||
169   TTGAGTCCAAACCCCGTAAAGACACCACTTATCGCACCTGGCAGCAGCCACTGGTAACAGA 110
      |||||
5806  TTAGCAGAGCGAGGTATGTAGCGGTCTACAGAGTTCCTGAAGTGGTGGCTTAACACG 5865
      |||||
109   TTACGAGCGAGGTATGTAGCGGTCTACAGAGTTCCTGAAGTGGTGGCTTAACACG 50
      |||||
5866  GCTACACTAGAAAGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGT 5914
      |||||
49   GSTACACTAGAAAGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGT 1

RESULT 15
CF269652
LOCUS
DEFINITION
    Fcylcold844_Fragilariopsis_cylindrus SMART cDNA library (Clontech)
    Fragilariopsis cylindrus cDNA clone Antarctic 5', mRNA sequence.
ACCESSION
CF269652
VERSION
CF269652.1
KEYWORDS
EST.
SOURCE
    Fragilariopsis cylindrus
    Fragilariopsis cylindrus
    Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
    Bacillariophycidae; Bacillariales; Bacillariaceae; Fragilariopsis.
REFERENCE
    1 (bases 1 to 1073)
    Mock, T. and Valentin, K.
    EST analysis of freezing tolerance in the Antarctic diatom
    Fragilariopsis cylindrus: Detection of numerous cold adaptation
    related genes and gene transfer events
    Unpublished (2003)
    Contact: Mock T
    Biological Oceanography
    Alfred-Wegener-Institute for Polar and Marine Research
    Am Handelshafen 12, D-27570 Bremerhaven, Germany
    Tel: +49 471 4831 1893
    Fax: +49 471 4831 1425
    Email: tmock@awi-bremerhaven.de
    sequence with unknown function
    PCR Primers
    FORWARD: 5'lambdaTriplex2

```



BACKWARD: 3'lambdaTriplex2  
Seq primer: ctcgggaagcgcgcattgtgttgg.  
Location/Qualifiers  
1. .1073  
/organism="Fragilariopsis cylindrus"  
/mol\_type="mRNA"  
/strain="Antarctic"  
/db\_xref="taxon:186039"  
/clone\_lib="Fragilariopsis cylindrus SMART cDNA library (Clontech)"  
/note="vector: pTriplex2; total polyA was used for first-strand synthesis with SMART IV oligos and CDS III/3'PCR primer. Double strand cDNA synthesis was done by LD PCR using the following program: 95OC for 5 min denaturation and subsequent 20 cycles at 95OC (2min) and 68OC (6min). After SfiI digestion the cDNA was fractionated with CHROMA Spin-400 columns. These cDNAs were ligated overnight into pTriplex2 vectors."

## ORIGIN

Query Match 11.3%; Score 865; DB 7; Length 1073;  
Best Local Similarity 96.2%; Pred. No. 1.9e-235;  
Matches 927; Conservative 0; Mismatches 30; Indels 7; Gaps 4;  
QY 5154 GCGTTCGCTCAGTCGCCGCTTCCAGTCGGGAACCTGTCGCCAGCTGCATTAATGA 5213  
DB 1 GCTCTTGTCTACATGCCCGCTTCCAGTCGGGAACCTGTCGCCAGCTGCATTAATGA 60  
QY 5214 ATCGGCCAAACCGCGGGGAGAGCGGTTTGGGTATTTGGGGCTCTTCGGCTTCCTCGCTC 5273  
DB 61 ATCGGCCAAACCGCGGGGAGAGCGGTTTGGGTATTTGGGGCTCTTCGGCTTCCTCGCTC 120  
QY 5274 ACTGACTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTATCAGCTCACTCAAGAGCG 5333  
DB 121 ACTGACTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTATCAGCTCACTCAAGAGCG 180  
QY 5334 GTATACGGTTATCCACAGAAATCAGGGGATTAACGAGGAAGAAACATGTGAGCAAAAGGC 5393  
DB 181 GTAATACGGTTATCCACAGAAATCAGGGGATTAACGAGGAAGAAACATGTGAGCAAAAGGC 240  
QY 5394 CAGCAAAAGCCAGGAACCGTAAAGCGCGCTGCTGGCGTTTTCATAGGCTCCGC 5453  
DB 241 CAGCAAAAGCCAGGAACCGTAAAGCGCGCTGCTGGCGTTTTCATAGGCTCCGC 300  
QY 5454 CCCCTCAGCAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGGAAACCCGACAGGA 5513  
DB 301 CCCCTCAGCAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGGAAACCCGACAGGA 360  
QY 5514 CTATAAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGGCTCTCCTGTTCCGACC 5573  
DB 361 CTATAAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGGCTCTCCTGTTCCGACC 420  
QY 5574 CTGGCGTTACCGGATACCTGTCCGCTTTCCTCCCTCGGGAAGCGTGGCGTTTCTCAT 5633  
DB 421 CTGGCGTTACCGGATACCTGTCCGCTTTCCTCCCTCGGGAAGCGTGGCGTTTCTCAT 480  
QY 5634 AGCTCAGCTGATAGTATCTCAGTTCGGTGTAGTTCGCTCCAGCTGGGCTGTGTG 5693  
DB 481 AGCTCAGCTGATAGTATCTCAGTTCGGTGTAGTTCGCTCCAGCTGGGCTGTGTG 540  
QY 5694 CAGGAACCCCGCTTACAGCCGACCGCTGGCGCTTATCCGGTAACTATCGTCTTGAGTCC 5753  
DB 541 CAGGAACCCCGCTTACAGCCGACCGCTGGCGCTTATCCGGTAACTATCGTCTTGAGTCC 600  
QY 5754 AACCCGTAAGACACGACTTATCGCACTGGCAGACGCACTGGTAACAGGATTAGCAGA 5813  
DB 601 AACCCGTAAGACACGACTTATCGCACTGGCAGACGCACTGGTAACAGGATTAGCAGA 660  
QY 5814 GCGAGGTATCTAGCGGTGTCTACAGAGTCTTGAAGTGGTGGCT--AACTACGGGTACA 5871  
DB 661 GCGAGGTATCTAGCGGTGTCTACAGAGTCTTGAAGTGGTGGCCNTAACTACNGGCTACA 720

Search completed: May 25, 2005, 19:41:54  
Job time : 14663 secs

QY 5872 CTAGAAGGACA-GTATTTGGTATCT-GCGCTCTGCTGAAGCC---AGTTACCTTTGGA 5926  
DB 721 CTAGAAGAACANGTATTTGGTATCTNGCGCTCTGCTNGAAGCCAGNTTACGNTTCGAAA 780  
QY 5927 AAGAGTTGGTAGCTCTTGATCCGGCAACAAACCCGCTGGTAGCGGTGTTTTTTTGT 5986  
DB 781 AAGAGTTGGTAGCTCTTGATCCGGCAACAAACCCGCTGGTAGCGGTGTTTTTTTGT 840  
QY 5987 TTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTC 6046  
DB 841 TTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTC 900  
QY 6047 TACGGGTCTTGACGCTCAGTGGAAACGAAACTCAAGTTAAGGATTTTGTGATGAGATT 6106  
DB 901 TACGGGTCTTGACGCTCAGTGGAAACGAAACTCAAGTTAAGGATTTTGTGATGAGATT 960  
QY 6107 ATCA 6110  
DB 961 ATAA 964

**THIS PAGE BLANK (USPTO)**

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	3111	40.6	5364	2	US-08-793-610-4	Sequence 4, Appli
2	3094	40.4	6365	1	US-08-352-990-1	Sequence 1, Appli
3	3083.4	40.3	5292	2	US-08-793-610-3	Sequence 3, Appli
4	3083.4	40.3	5294	2	US-08-793-610-2	Sequence 2, Appli
5	3083.4	40.3	5323	2	US-08-793-610-1	Sequence 1, Appli
6	3062.2	40.0	6145	1	US-08-336-132-1	Sequence 1, Appli
7	3062.2	40.0	6145	3	US-08-935-312-1	Sequence 1, Appli
8	3062.2	40.0	6145	3	US-08-848-760B-1	Sequence 1, Appli
9	3062.2	40.0	6145	4	US-09-826-025-1	Sequence 1, Appli
10	3050.2	39.9	6620	4	US-08-786-531B-3	Sequence 3, Appli
11	2703	35.3	5865	3	US-08-654-737B-3	Sequence 3, Appli
12	2672.4	34.9	7086	4	US-09-935-194-1	Sequence 1, Appli
13	2633.2	34.4	7560	3	US-08-844-274-20	Sequence 20, Appli
14	2633.2	34.4	7560	4	US-09-598-421-20	Sequence 20, Appli
15	2623.4	34.3	9164	4	US-09-205-817A-2	Sequence 2, Appli
16	2623.4	34.3	11846	4	US-09-205-817A-4	Sequence 4, Appli
17	2623.4	34.3	12022	4	US-09-205-817A-3	Sequence 3, Appli
18	2623.2	34.3	7160	4	US-08-786-531B-5	Sequence 5, Appli
19	2623	34.3	4016	3	US-09-173-053-3	Sequence 3, Appli
20	2621.2	34.2	4950	3	US-08-789-333F-58	Sequence 58, Appli
21	2621.2	34.2	4950	3	US-08-589-109A-7	Sequence 7, Appli
22	2621.2	34.2	4950	3	US-08-787-738B-58	Sequence 58, Appli
23	2621.2	34.2	4950	4	US-09-916-940-58	Sequence 58, Appli
24	2621.2	34.2	4950	4	US-10-096-550-58	Sequence 58, Appli
25	2620	34.2	9639	3	US-09-147-208-26	Sequence 26, Appli
26	2620	34.2	9639	4	US-09-550-117A-26	Sequence 26, Appli
27	2618	34.2	9632	4	US-09-403-752A-1	Sequence 1, Appli

Query Match 40.6%; Score 3111; DB 2; Length 5364;									
Best Local Similarity 92.0%; Pred. No. 0; Mismatches 210; Indels 87; Gaps 9;									
Matches 3414; Conservative 0;									
QY	3868	ACGAGCCATAGATAAATAAAGATTTTATTTAGTCTCCAGAAAAGGGGGGAATCAAG	3927						
DB	1655	ACGAGCCATAGATAGATAAAGATTTTATTTAGTCTCCAGAAAAGGGGGGAATCAAG	1714						
QY	3928	ACCCACCTGTAGTTTGGCAAGC-----TAGCTTAAGTAAAGCCAT	3969						
DB	1715	ACCCACCTGTAGTTTGGCAAGC-----TAGCTTAAGTAAAGCCAT	1774						
QY	3970	TTTGAAGGCATGG-AAAATACATACATCAGATAAGAGTTTCAGATCAAGTTAGGAA	4028						
DB	1775	TTTGAAGGCATGGAAAATAACAAACCAAGAAATAGGAAGTTTCAGATCAAGGGCGGTA	1834						
QY	4029	CAGAGAGACGACAGATATGGGCAAAACAGGATATCTGTGTAAGCAGTTCTTGCCTCCG	4088						
DB	1835	CATGAAAATAGCTAACGTTGGGCCAAACAGGATATCTGCGGTGAGCAGTTTCGGGCCCGG	1894						
QY	4089	TCAGGGCCAAAGACAGTTGGAAACAGGAAATATGGGCCAAAC-----AGGATATCTG	4140						
DB	1895	CCCGGGGCAAGAACAGATGGTCAACGACGTTTCGGGCCCGGCCCGGCCCAAGAACAGATG	1954						
QY	4141	TGGTAAGCAGTTCTCCGCCCGCTCAGGGCCAAAGAACAGATGGTCCCAAGATCGGTCCC	4200						
DB	1955	GTCACGCGAGTTTCGGGCCCGGCCCGGCCCGGCCCAAGAACAGATGGTCCCAAGATATG	2014						
QY	4201	GCCTCAGCAGTTCTCAGAGAACCATCAGATGTTTCCAGGGTGCCCCCAAGGACCTGAAT	4260						
DB	2015	ACCCTCAGCAGTTCTTAAAGACCCATCAGATGTTTCCAGGCTCCGCCCAAGGACCTGAAT	2074						
QY	4261	GACCTGTGCTTATTTGAATCAACCATCAGTTGCTTCTCGCTTCTGTTTCGGGGCTT	4320						
DB	2075	GACCTGTGCTTATTTGAATCAACCATCAGCTGCTTCTCGCTTCTGTTTCGGGGCTT	2134						
QY	4321	CTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGGCGCCAGTCTCCGA	4380						
DB	2135	CTGCTTCCCGAGCTCATAAAGAGCTCACACCCCTCACTCGGGCGCCAGTCTCCGA	2194						
QY	4381	TAGACTGCTGCGCCCGGTACCGGTGTTCTCAATAAACCCCTCTTGAGTTGATCCGACT	4440						
DB	2195	TTGACTGAGTCCGCCGGGTACCGGTGTTCTCAATAAACCCCTCTTGAGTTGATCCGACT	2254						
QY	4441	CGTGTCTCGCTGCTTCTTGGAGGGTCT-CTCTGAGTGAATTGACTACCCGTCAGC-GGG	4498						
DB	2255	CGTGTCTCGCTGCTTCTTGGAGGGTCTCTCTGAGTGAATTGACTACCCGTCAGCGGG	2314						
QY	4499	GTCTTTCAGTTCTCCACCTACACAGGTCTCACTAAACATTCCTGATGTGCCGAGGGAC	4558						
DB	2315	GTCTTTCAGTTCTCTCCACCTACACAGGTCTCACTAAACATTCCTGATGTGCCGAGGGAC	2374						
QY	4559	TCGTCAGCCCGGTTTTTTGTTTATATAAATAAAGCAAGAACAGTGTTCCTTCAAGCCAGA	4618						
DB	2375	TCGTCAGCCCGGTTTTTTGTTTATATAAATAAAGCAAGAACAGTGTTCCTTCAAGCCAGA	2434						
QY	4619	CTACATCTCTGCTCTCGGCTTTATAAAGAAATGTTGAAGGGCTCTGTGAGCTATCTGCCA	4678						
DB	2435	CTACATCTCTGCTCTCGGCTTTATAAAGAAATGTTGAAGGGCTCTGTGAGCTATCTGCCA	2494						
QY	4679	CACGACTTTTAAAGATTTTATGCTCTCCGATGAGGGATTTAGTCAATCTATCTCTCGTC	4738						
DB	2495	CACGAC-TTTTAAAGATTTTATGCTCTCCGATGAGGGATTTAGTCAATCTATCTCTCGTC	2553						
QY	4739	TATTTTGTGCTTCTCCGTAATTTTAAATTTCTAGTTTGCACCTCCCTCTCAGAGCACG	4798						
DB	2554	TATTTTGTGCTTCTCCGTAATTTTAAATTTCTAGTTTGCACCTCCCTCTCAGAGCACG	2613						
QY	4799	GGGATTGCAAGTAGTTAATCTCTGAGGGCGAGCTTCTGTGAAGAGTTTGGCTGGGCTC	4858						
DB	2614	GGGATTGCAAGTAGTTAATCTCTGAGGGCGAGCTTCTGTGAAGAGTTTGGCTGGGCTC	2673						
QY	4859	AGTGTGAGATTTTGCCATATAAAGGGGTCTGCCCTGTGTACAGACAGATCGGAATCTA	4918						
DB	2674	AGTGTGAGATTTTGCCATATAAAGGGGTCTGCCCTGTGTACAGACAGATCGGAATCTA	2733						
QY	4919	GAGTGCATACCTCAGAGTCCCCCGGTTTCGGGGCTCTGATCTCAGGGCAATCTTTGCTTAG	4978						
DB	2734	GAGTGCATACCTCAGAGTCCCCCGGTTTCGGGGCTCTGATCTCAGGGCAATCTTTGCTTAG	2793						
QY	4979	AGATCCTCTACGCCGGAACATCGTGGCGGGGTACCGAGCTCGAATTCGTATCATGGTC	5038						
DB	2794	AGATCCTCTACGCCGGAACATCGTGGCGGGGTACCGAGCTCGAATTCGTATCATGGTC	2853						
QY	5039	ATAGCTGTTC-----CTGTGTGAAATTTGTTATCCGCTCAAAATCCACACA	5085						
DB	2854	GGGCTTATATCGCCGACATCACCGATGGGAAGATCGGGCTCGCCACTTCGGGCTCATGA	2913						
QY	5086	ACATACGAGCCGGAAGCAATAAGTGTAAAGCTTGGGGTGCCTTAATGAAGT-----	5134						
DB	2914	GGGCTTGTTCGGCGTGGGTATGTTGGCAGGCCCCGTGGCCGGGGAATGTTGGGCGCCA	2973						
QY	5135	-----GAGCTAACTCACATTAATTCGGTTGCG	5162						
DB	2974	TCTCTTCGATGACCATTCCTTTCGGGGCGGTGCTCAACGGCTCAACCTACTACTCG	3033						
QY	5163	TCACCTGCCGCTTTCAGTCCGGGAAACCTGTCTGCCAGCTGCATTAATGAATCGGCCAA	5222						
DB	3034	GCTGCTTCTTAATGACGAGTCCGATAGGGAGAGGCTCTGCAATTAATGAATCGGCCAA	3093						
QY	5223	CGCGGGGAGAGCGGTTTGGTATTTGGCGCTCTTCGCTTCCTCGCTCACTGACTCG	5282						
DB	3094	CGCGGGGAGAGCGGTTTGGTATTTGGCGCTCTTCGCTTCCTCGCTCACTGACTCG	3153						
QY	5283	CTGCGCTCGTCTTCGCTCGCGGCGGGTATCAGCTCACTCAAGGCGGTATACGG	5342						
DB	3154	CTGCGCTCGTCTTCGCTCGCGGCGGGTATCAGCTCACTCAAGGCGGTATACGG	3213						
QY	5343	TTATCCACAGATCAGGGGATAACGACAGAAAGAAACATGTAGCAAAAGGCCAGCAAAAG	5402						
DB	3214	TTATCCACAGATCAGGGGATAACGACAGAAAGAAACATGTAGCAAAAGGCCAGCAAAAG	3273						
QY	5403	GCCAGAAACCGTAAAGGCCGCTTGTGTGGCTTTTCCATAGGCTCGGCCCTCGAC	5462						
DB	3274	GCCAGAAACCGTAAAGGCCGCTTGTGTGGCTTTTCCATAGGCTCGGCCCTCGAC	3333						
QY	5463	GAGCATCAAAATCGAGCTCAAGTCAGAGTGGCGAACCACGACAGGACTATAAGA	5522						
DB	3334	GAGCATCAAAATCGAGCTCAAGTCAGAGTGGCGAACCACGACAGGACTATAAGA	3393						
QY	5523	TACAGGGGTTTCCCTTGGAGCTCCCTCGTGCGCTCTCTCTGTTCGACCCCTGCGCTT	5582						
DB	3394	TACAGGGGTTTCCCTTGGAGCTCCCTCGTGCGCTCTCTCTGTTCGACCCCTGCGCTT	3453						
QY	5583	ACCGATACCTGTTCGCTTTCCTTCGGGAAAGGCGGCGCTTCTCATAGTCAAGC	5642						
DB	3454	ACCGATACCTGTTCGCTTTCCTTCGGGAAAGGCGGCGCTTCTCATAGTCAAGC	3513						
QY	5643	TGTAGGTATCTCAGTTTCGGTGTAGTCTCGCTCAAGCTGGGCTGTGACGCAACCC	5702						
DB	3514	TGTAGGTATCTCAGTTTCGGTGTAGTCTCGCTCAAGCTGGGCTGTGACGCAACCC	3573						
QY	5703	CCCGTTACGCGCCGCTGCGCTTATCCGGTAACTATCGCTTTCGAGTCCAAACCCGTA	5762						
DB	3574	CCCGTTACGCGCCGCTGCGCTTATCCGGTAACTATCGCTTTCGAGTCCAAACCCGTA	3633						
QY	5763	AGACAGCATTTATCCGCACTGGCAGCAGCACTGTGTAAACAGATTAGCAGAGCGAGTAT	5822						
DB	3634	AGACAGCATTTATCCGCACTGGCAGCAGCACTGTGTAAACAGATTAGCAGAGCGAGTAT	3693						
QY	5823	GTAGCGGCTGTACAGAGTTCTTAAGTGGTGGCTTAAGTGGCTTACCTAGAGAGGACA	5882						
DB	3694	GTAGCGGCTGTACAGAGTTCTTAAGTGGTGGCTTAACTAGCGGTACACTAGAGAGGACA	3753						
QY	5883	GTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACTCTTCGGAAGAGGTTGCTGCTCT	5942						









Db 3982 CCCTCAGCGGGGCTTTTCAATTTGGGGCTCGTCCGGATCGGGAGACCCCTGCCAGG 4041  
 QY 4547 TGCCGAGGAGCTCCGCTCAGCCCGTTTTTGTATTATAAATAAGCAAGATGTTCC 4606  
 Db 4042 GACCAACCGACCAACCGGAGTAAGCTG----- 4072  
 QY 4607 CTTCAAGCCAGACTACATCCTCGACTCTCGGCTTTATAAAGAAATGTTGAAGGCTCTGTG 4666  
 Db 4073 -----GCTGCTCGCGGTTTCGGTGATGACGGTGAACCTCTG-A 4113  
 QY 4667 GACTATCTGCGCACAGACTTTTAAAGATTTTATGCTCCTCGATGAGGATTTAGTCAA 4726  
 Db 4114 CACATCAGCTCCCGAGACGGTCAAGCTTGTCTGTAAGCGGATCCCGGAGACACAA 4173  
 QY 4727 TCTATCCTCGTCTATTTTCTGCTTCTCGTCTCTCGTATTTTAAATTTCTAGTTTGCACCTCT 4786  
 Db 4174 GCCCGTCAGGCGCGTCAGCGGTGTTGGC----- 4204  
 QY 4787 CCTGAGAGCAGCGGATTCAGAGTAGTTAATACTCTGAGGGCAGGCTTCTGTGAAAAGG 4846  
 Db 4205 ----- 4204  
 QY 4847 TTGCTGGGCTCAGTGTGAGATTTTCCCATAAAGGGTCTCTGCCCTGTGTACAGACA 4906  
 Db 4205 ----- 4204  
 QY 4907 GATCGGAATCAGATGTCATCTCAGAGTCCCGCGGTTCCGGGCTCTGATCTCAGGGC 4966  
 Db 4205 -----GGTGTGGGGCGCAGCCATGACCCA 4229  
 QY 4967 ATCTTTTCCCTAGAGATCCTCTAGCCCGAGCGATCGTGCCGGTACCGAGCTCGAATTC 5026  
 Db 4230 GTC-----ACGTAGCATAGCGGAGTGATATCTGGCTTAA 4264  
 QY 5027 GTAATCATGTCATAGCTGTTCTGTGTGAATTTGTTATCCGCTCACAAATTCACACAA 5086  
 Db 4265 CTATGGGGCATCAGACAGATTGTAAGAG----- 4296  
 QY 5087 CATACGAGCGGAGCATAAAGTGTAAAGCTGGGGTGCTTAATGAGTGAGCTAATCTAC 5146  
 Db 4297 ----- 4296  
 QY 5147 ATTAATTTGGCTTTCGCTCACTGCCGCTTTTCCAGTCGGGAACCTGCTGCGCAGTGCA 5206  
 Db 4297 -----TGACCATATGGC 4309  
 QY 5207 TTAATGAATCGGCACACGCGCGGGAGAGCGGTTTTCGTAATGGCGCTCTTCCGCTTC 5266  
 Db 4310 GTGTGAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCTCTTCCGCTTC 4369  
 QY 5267 CTGCTCACTGACTCGCTCGCTCGTTCGCTGCGGCTGCGGAGCGGATCAGCTCCTC 5326  
 Db 4370 CTGCTCACTGACTCGCTCGCTCGTTCGCTGCGGCTGCGGAGCGGATCAGCTCCTC 4429  
 QY 5327 AAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGAGGAAGAAATGAGGAC 5386  
 Db 4430 AAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGAGGAAGAAATGAGGAC 4489  
 QY 5387 AAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGAGGAAGAAATGAGGAC 5446  
 Db 4490 AAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGAGGAAGAAATGAGGAC 4549  
 QY 5447 GCTCCGCCCTTGACGAGATCACAATAATGACGCTCAAGTCAAGAGTGGCGAAACCC 5506  
 Db 4550 GCTCCGCCCTTGACGAGATCACAATAATGACGCTCAAGTCAAGAGTGGCGAAACCC 4609  
 QY 5507 GACGAGCTATAAAGATACAGCGGTTTCCCTGGAAGCTCCCTGCGGCTCTCCTGT 5566  
 Db 4610 GACGAGCTATAAAGATACAGCGGTTTCCCTGGAAGCTCCCTGCGGCTCTCCTGT 4669  
 QY 5567 TCCGACCTTCCGCTTTACCGGATACCTGTCCGCTTTTCTCCCTTCGCGGAGCGTGGCGCT 5626  
 Db 4670 TCCGACCTTCCGCTTTACCGGATACCTGTCCGCTTTTCTCCCTTCGCGGAGCGTGGCGCT 4729

QY 5627 TTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGTTAGGTGCTTCCGTTCCAGCTGG 5686  
 Db 4730 TTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGTTAGGTGCTTCCGTTCCAGCTGG 4789  
 QY 5687 CTGTGTGACGAAACCCCGTTTACCGGATCGGCTTATCGGTTAACTATCGTCT 5746  
 Db 4790 CTGTGTGACGAAACCCCGTTTACCGGATCGGCTTATCGGTTAACTATCGTCT 4849  
 QY 5747 TGAGTCCAAACCCCGTTTACCGGATCGGCTTATCGGTTAACTATCGGTTAACTATCGG 5806  
 Db 4850 TGAGTCCAAACCCCGTTTACCGGATCGGCTTATCGGTTAACTATCGGTTAACTATCGG 4909  
 QY 5807 TAGCAGAGCGAGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGTGTCCTAACTACCG 5866  
 Db 4910 TAGCAGAGCGAGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGTGTCCTAACTACCG 4969  
 QY 5867 CTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAA 5926  
 Db 4970 CTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAA 5029  
 QY 5927 AAGAGTTGGTAGTCTTGTATCCGGCAAAACCAACCGCTGCTGAGGGTGGTTTTTGT 5986  
 Db 5030 AAGAGTTGGTAGTCTTGTATCCGGCAAAACCAACCGCTGCTGAGGGTGGTTTTTGT 5089  
 QY 5987 TTGCAAGCAGCAGATTTAGCGCGCAGAAAAAGGATCTCAAGAGATCTTTCATCTTTTC 6046  
 Db 5090 TTGCAAGCAGCAGATTTAGCGCGCAGAAAAAGGATCTCAAGAGATCTTTCATCTTTTC 5149  
 QY 6047 TACGGGCTCTGACGCTCAGTGAACGAAACCTCAGTTAAGGATTTTGGTCATGAGATT 6106  
 Db 5150 TACGGGCTCTGACGCTCAGTGAACGAAACCTCAGTTAAGGATTTTGGTCATGAGATT 5209  
 QY 6107 ATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATAAAGGATTTTAAATCAATCTA 6166  
 Db 5210 ATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATAAAGGATTTTAAATCAATCTA 5269  
 QY 6167 AAGTATATAGTAAACCTTGTCTGACAGTTTACCAATGCTTAATCAGTGAGGACCTAT 6226  
 Db 5270 AAGTATATAGTAAACCTTGTCTGACAGTTTACCAATGCTTAATCAGTGAGGACCTAT 5329  
 QY 6227 CTCAGGATCTGCTATTTTCTCATCCATAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6286  
 Db 5330 CTCAGGATCTGCTATTTTCTCATCCATAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5389  
 QY 6287 TACGATAGCGGAGGCTTACCATCTGCGCCCAAGTCTGCAATGATACCGCGAGACCCAG 6346  
 Db 5390 TACGATAGCGGAGGCTTACCATCTGCGCCCAAGTCTGCAATGATACCGCGAGACCCAG 5449  
 QY 6347 CTCAGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGCGGAGGAGAG 6406  
 Db 5450 CTCAGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGCGGAGGAGAG 5509  
 QY 6407 TGGTCTGCAACTTTTATCGGCTCCATCCAGTCTAATTAATTTGTCGGGAAAGCTAGAGT 6466  
 Db 5510 TGGTCTGCAACTTTTATCGGCTCCATCCAGTCTAATTAATTTGTCGGGAAAGCTAGAGT 5569  
 QY 6467 AAGTATAGTTCGCGATTAATGTTTCGCAACGTTGTCCTATGCTTACAGGC-TGCTGT 6525  
 Db 5570 AAGTATAGTTCGCGATTAATGTTTCGCAACGTTGTCCTATGCTTACAGGCATGCTGT 5629  
 QY 6526 GTACGCTCGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 6585  
 Db 5630 GTACGCTCGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 5689  
 QY 6586 TACATGATCCCGATCTGTTGTCAAAAAGCGTTAGCTCTTTCGCTCTCCGATCGTTGT 6645  
 Db 5690 TACATGATCCCGATCTGTTGTCAAAAAGCGTTAGCTCTTTCGCTCTCCGATCGTTGT 5749  
 QY 6646 CAGAGTAAAGTTGGCGCAGTGTATCACTCATGTTTGGCAGCACTGTCATAAATCTCT 6705  
 Db 5750 CAGAGTAAAGTTGGCGCAGTGTATCACTCATGTTTGGCAGCACTGTCATAAATCTCT 5809



Db 2311 CCGGTTTGTGTTTAAATAAATGCAAGACAGTGTCCCTTCAAGCCAGACTACATCC 2370  
QY  
QY 4627 TGACTCTCGGCTTTATAAAGAAATGTTGAAGGCTCTGTGGACTATCTGCCACACGACTT 4686  
Db  
Db 2371 TGACTCTCGGCTTTATAAAGAAATGTTGAAGGCTCTGTGGACTATCTGCCACACGAC-T 2429  
QY 4687 TTTAAGATTTTATGCTCTCTGATGAGGGAATTTAGTCAATCTATCTCTGCTATTTTGC 4746  
Db  
Db 2430 TTTAAGATTTTATGCTCTCTGATGAGGGAATTTAGTCAATCTATCTCTGCTATTTTGC 2489  
QY 4747 TGGCTTCTCGGATTTTAAATTTCTAGTTTGCACCTCCCTCTGAGAGCACGGCGATTC 4806  
Db  
Db 2490 TGGCTTCTCGGATTTTAAATTTCTAGTTTGCACCTCCCTCTGAGAGCACGGCGATTC 2549  
QY 4807 AGAGTAGTTAACTACTCTGAGGCGAGGCTTCTGTGAAGAGTTGCTGGGCTCAGTGTGAG 4866  
Db  
Db 2550 AGAGTAGTTAACTACTCTGAGGCGAGGCTTCTGTGAAGAGTTGCTGGGCTCAGTGTGAG 2609  
QY 4867 ATTTTGGCCATAAAGAGGCTCTGCCCTGTGTGACAGAGATCGGAATCTAGAGTGCAT 4926  
Db  
Db 2610 ATTTTGGCCATAAAGAGGCTCTGCCCTGTGTGACAGAGATCGGAATCTAGAGTGCAT 2669  
QY 4927 ACTCAGAGTCCCGGCTTCCGGGCTCTGATCTCAGGGCATCTTTGCCCTAGAGATCTTC 4986  
Db  
Db 2670 ACTCAGAGTCCCGGCTTCCGGGCTCTGATCTCAGGGCATCTTTGCCCTAGAGATCTTC 2729  
QY 4987 TAGCCGAGCGCATCTGCGCGGCTACCGAGCTCGAATTCGTAATCATGTCATAGCTGT 5046  
Db  
Db 2730 TAGCCGAGCGCATCTGCGCGGCTACCGAGCTCGAATTCGTAATCATGTCATAGCTGT 2789  
QY 5047 TTC-----CTGTGTGAATTTGTATTCGCTCACAAATTCACACAACTACGA 5093  
Db  
Db 2790 ATCCGCGCATCACCGATGGGAGATCGGCTCGCCACTTCGGGCTCATGAGCGTGT 2849  
QY 5094 GCGGAAGCATAAAGTGAAGCTTGGGCTGCTTAATGAGT----- 5134  
Db  
Db 2850 TTGCGGCTGGTATGGTGAGCGCCCGCTGGCGGGGACTGTGGCGGCCACTCTCCTTG 2909  
QY 5135 -----GAGCTAACTCACAATTAATGCGTTCGCTCACTGCC 5170  
Db  
Db 2910 CATGCAACATTCCTCGCGGCGGCTCAACGGCTCAACCTACTCTGCGGCTGCTTC 2969  
QY 5171 CGCTTTCCAGTGGGAAACCTGTCTGCCAGCTGCAATTAATGAATCGGCAACGGCGGG 5230  
Db  
Db 2970 CTAACTCAGAGTGCATGAAGGAGAGCGTCTGCAATTAATGAATCGGCAACGGCGGG 3029  
QY 5231 GAGAGCGGTTTGGGCTTTCGCTTCTCGCTCACTGACTCGCTGCGCTC 5290  
Db  
Db 3030 GAGAGCGGTTTGGGCTTTCGCTTCTCGCTCACTGACTCGCTGCGCTC 3089  
QY 5291 GGTCTGCTGCGGCGGCTATCAGCTCACTCAAGGCGGTAATACGGTTATCCAC 5350  
Db  
Db 3090 GGTCTGCTGCGGCGGCTATCAGCTCACTCAAGGCGGTAATACGGTTATCCAC 3149  
QY 5351 AGAATCAGGGATAACGAGGAAAGAAATGTGAGCAAAAGGCAGCAAAAGCCAGGAA 5410  
Db  
Db 3150 AGAATCAGGGATAACGAGGAAAGAAATGTGAGCAAAAGGCAGCAAAAGCCAGGAA 3209  
QY 5411 CCGTAAAAAGCGGCTTCTGCGGTTTTTCCATAGGCTCCGCCCTCCCTGACGAGCATCA 5470  
Db  
Db 3210 CCGTAAAAAGCGGCTTCTGCGGTTTTTCCATAGGCTCCGCCCTCCCTGACGAGCATCA 3269  
QY 5471 CAAAAATCAGCGTCAAGTCAGAGTGGGAAACCGGACAGACTATATAAGATACCGAG 5530  
Db  
Db 3270 CAAAAATCAGCGTCAAGTCAGAGTGGGAAACCGGACAGACTATATAAGATACCGAG 3329  
QY 5531 GTTTCCTCCCTGGAAGCTCCCTGTGCGCTCTCTGTTCCGACCCCTGCGCTTACCGGATA 5590  
Db  
Db 3330 GTTTCCTCCCTGGAAGCTCCCTGTGCGCTCTCTGTTCCGACCCCTGCGCTTACCGGATA 3389  
QY 5591 CTGTCCGCTTTCCTCCCTTCCGGAAGCGTGGCGCTTCTCATAGCTCACTGTAGGTA 5650  
Db  
Db 3390 CTGTCCGCTTTCCTCCCTTCCGGAAGCGTGGCGCTTCTCATAGCTCACTGTAGGTA 3449

QY 5651 TCTCAGTTTGGTGTAGTTCGTTCCGCTCCAAGCTGGGCTGTGTGACGAACCCCGGTTCA 5710  
Db  
Db 3450 TCTCAGTTTGGTGTAGTTCGTTCCGCTCCAAGCTGGGCTGTGTGACGAACCCCGGTTCA 3509  
QY 5711 GCCGACCGCTGCGGCTTATCCGGTAACTATCGTCTTGAGTCCAAACCGGTTAAGACAGA 5770  
Db  
Db 3510 GCCGACCGCTGCGGCTTATCCGGTAACTATCGTCTTGAGTCCAAACCGGTTAAGACAGA 3569  
QY 5771 CTTATCGCCACCTGCGAGCAGCACCTGTAACAGGATTAAGCAGAGCGAGTATGTTGGCG 5830  
Db  
Db 3570 CTTATCGCCACCTGCGAGCAGCACCTGTAACAGGATTAAGCAGAGCGAGTATGTTGGCG 3629  
QY 5831 TGCTACAGATTTCTTGAAGTGGTGCCTAACTTACGGCTACACTAGAAGCAGATTTTGG 5890  
Db  
Db 3630 TGCTACAGATTTCTTGAAGTGGTGCCTAACTTACGGCTACACTAGAAGCAGATTTTGG 3689  
QY 5891 TATCTGCGCTCTGTGAAGCCAGTTTACCTTCGGAAGAAAGAGTTGTAGCTTTGTATTCGG 5950  
Db  
Db 3690 TATCTGCGCTCTGTGAAGCCAGTTTACCTTCGGAAGAAAGAGTTGTAGCTTTGTATTCGG 3749  
QY 5951 CAAACAAACCCCGCTGCTAGCGGTGTTTTTTGTTTGAAGCAGCAGATTAACGGCAG 6010  
Db  
Db 3750 CAAACAAACCCCGCTGCTAGCGGTGTTTTTTGTTTGAAGCAGCAGATTAACGGCAG 3809  
QY 6011 AAAAAAGGATCTCAAGAGATCTCTTCTGATCTTTTCTACGGGCTCTGACGCTCAGTGGAA 6070  
Db  
Db 3810 AAAAAAGGATCTCAAGAGATCTCTTCTGATCTTTTCTACGGGCTCTGACGCTCAGTGGAA 3869  
QY 6071 CGAAAACTCAGCTTAAGGATTTTGTCTAGATTAATCAAAAAAGATCTTCACTTAGAT 6130  
Db  
Db 3870 CGAAAACTCAGCTTAAGGATTTTGTCTAGATTAATCAAAAAAGATCTTCACTTAGAT 3929  
QY 6131 CTTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATGATGATAAAGTTGTC 6190  
Db  
Db 3930 CTTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATGATGATAAAGTTGTC 3989  
QY 6191 TGACAGTTTACCAATCTTAATCAGTGAGGCACTATCTCAGCGATCTGTCTATTTTCGTTTC 6250  
Db  
Db 3990 TGACAGTTTACCAATCTTAATCAGTGAGGCACTATCTCAGCGATCTGTCTATTTTCGTTTC 4049  
QY 6251 ATCCATAGTTGCTGACTCCCGCTCGTGTAGATAAATCAAGATACGGGAGGCTTACCATC 6310  
Db  
Db 4050 ATCCATAGTTGCTGACTCCCGCTCGTGTAGATAAATCAAGATACGGGAGGCTTACCATC 4109  
QY 6311 TGGCCCCAGTGTGCAATGATACCGGAGACCCAGCTCACCGCTCCAGATTTATCAGC 6370  
Db  
Db 4110 TGGCCCCAGTGTGCAATGATACCGGAGACCCAGCTCACCGCTCCAGATTTATCAGC 4169  
QY 6371 AATAAACCCAGCCAGCGGAAGCGGCGAGCGAGAGTGTCTGCAACTTTTATCCGCTTC 6430  
Db  
Db 4170 AATAAACCCAGCCAGCGGAAGCGGCGAGCGAGAGTGTCTGCAACTTTTATCCGCTTC 4229  
QY 6431 CATCCAGTCTATTAATTTGTCGGGGAAGCTAGAGTAAGTGTTCGCCAGTTAATAGTTT 6490  
Db  
Db 4230 CATCCAGTCTATTAATTTGTCGGGGAAGCTAGAGTAAGTGTTCGCCAGTTAATAGTTT 4289  
QY 6491 GCGCAACGTTGTGCCATGTCTACAGGC-TGTTGTGTCAAGCTCCGCTCGTGTGTTGTTATGCG 6549  
Db  
Db 4290 GCGCAACGTTGTGCCATGTCTACAGGCATCGTGTGTCAAGCTCCGCTCGTGTGTTGTTATGCG 4349  
QY 6550 TTCAATTCAGCTCCGCTTCCCAACGATCAAGGCGAGTTACATGATCCCATGTTGTGCAA 6609  
Db  
Db 4350 TTCAATTCAGCTCCGCTTCCCAACGATCAAGGCGAGTTACATGATCCCATGTTGTGCAA 4409  
QY 6610 AAAAGCGTTAGCTCTCTTCCGTTCCCGATCGTTGTGAGAGTAAGTTGGCCGCGAGTGT 6669  
Db  
Db 4410 AAAAGCGTTAGCTCTCTTCCGTTCCCGATCGTTGTGAGAGTAAGTTGGCCGCGAGTGT 4469  
QY 6670 ATCACTCATGTTTATGGCAGCATGCTGATAATTTCTTTACTGTCTATGCGCATCCGTAAGATG 6729  
Db  
Db 4470 ATCACTCATGTTTATGGCAGCATGCTGATAATTTCTTTACTGTCTATGCGCATCCGTAAGATG 4529



```

QY 6730 CTTTCTGAGTGGTGTGAGTACTCAACCAAGTCAATCTGAGAAATAGTGTATCGCGGACC 6789
Db 4530 CTTTCTGAGTGGTGTGAGTACTCAACCAAGTCAATCTGAGAAATAGTGTATCGCGGACC 4589
QY 6790 GAGTTGCTCTTCCCGCGGTCAATACCGGATAATACCGGCACATAGCAGAACTTTAAA 6849
Db 4590 GAGTTGCTCTTCCCGCGGTCAATACCGGATAATACCGGCACATAGCAGAACTTTAAA 4649
QY 6850 AGTGCTCATCATTTGAAAAAGTTTCTCGGGGCGAAACTCTCAAGATCTTTACCGCTGT 6909
Db 4650 AGTGCTCATCATTTGAAAAAGTTTCTCGGGGCGAAACTCTCAAGATCTTTACCGCTGT 4709
QY 6910 GAGATCCAGTTCGATGTAACCACTCGTCACCACTGATCTTCAGCATCTTTACTTT 6969
Db 4710 GAGATCCAGTTCGATGTAACCACTCGTCACCACTGATCTTCAGCATCTTTACTTT 4769
QY 6970 CACCAAGCTTTCTGGGTGAGCAAAACAGGAAGGCAAAATCGCAAAAAGGGAATAG 7029
Db 4770 CACCAAGCTTTCTGGGTGAGCAAAACAGGAAGGCAAAATCGCAAAAAGGGAATAG 4829
QY 7030 GCGCACGGAATGTTGAATCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7089
Db 4830 GCGCACGGAATGTTGAATCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4889
QY 7090 TCAGGGTTATGTTCTCATGAGCGGATACATATTGAAATGTTTGAATAAATAAATAAT 7149
Db 4890 TCAGGGTTATGTTCTCATGAGCGGATACATATTGAAATGTTTGAATAAATAAATAAT 4949
QY 7150 AGGGGTTCCGCGCACATTTCCCGGAAAGTGCACCTGACGCTCAAGAAACCATATTAT 7209
Db 4950 AGGGGTTCCGCGCACATTTCCCGGAAAGTGCACCTGACGCTCAAGAAACCATATTAT 5009
QY 7210 CATGACATTAACCTTAAAAATAGCGGTATCAGAGGCGCTTCTCTCGCGGTTTCGG 7269
Db 5010 CATGACATTAACCTTAAAAATAGCGGTATCAGAGGCGCTTCTCTCGCGGTTTCGG 5069
QY 7270 TGATGACGGTGAAACCTCTGACACATGACGCTCCCGAGACGCTCAAGCTTGTCTGTA 7329
Db 5070 TGATGACGGTGAAACCTCTGACACATGACGCTCCCGAGACGCTCAAGCTTGTCTGTA 5129
QY 7330 AGCGATCCGGGACGACAGACGCGCTCAGGCGCGCTCAGGCGGTGTTGGGGGTGCG 7389
Db 5130 AGCGATCCGGGACGACAGACGCGCTCAGGCGCGCTCAGGCGGTGTTGGGGGTGCG 5189
QY 7390 GGGCTGGTTAACTATGCGGCATCAGAGCAGATTTGCTGAGTGACACCATATGCGGTG 7449
Db 5190 GGGCTGGTTAACTATGCGGCATCAGAGCAGATTTGCTGAGTGACACCATATGCGGTG 5249
QY 7450 TGAATACCGCACAGATCGTAAGAGAAATAACCGCATCAGG 7492
Db 5250 TGAATACCGCACAGATCGTAAGAGAAATAACCGCATCAGG 5292

```

```

RESULT 4
US-08-793-610-2
; Sequence 2, Application US/08793610
; Patent No. 5858744
; GENERAL INFORMATION:
; APPLICANT: BAUM, Christopher
; APPLICANT: STOCKING-HARBERS, Carol
; APPLICANT: OSTERTAG, Wolfram
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF
; TITLE OF INVENTION: FOR GENE TRANSFER
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,610
; FILING DATE: 07-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 31 973.8
; FILING DATE: 08-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 195 03 952.1
; FILING DATE: 07-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/03175
; FILING DATE: 10-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Berman, Richard J.
; REGISTRATION NUMBER: 39,105
; REFERENCE/DOCKET NUMBER: P1614-7007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5294 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA
US-08-793-610-2

Query Match 40.3%; Score 3083.4; DB 2; Length 5294;
Best Local Similarity 91.8%; Pred. No. 0;
Matches 3400; Conservative 0; Mismatches 191; Indels 112; Gaps 9;

QY 3868 ACAGCCATAGATAAAATAAAAGATTTTATTAGTCTCCAGAAAAGGGGGGAATGAAG 3927
Db 1626 ACAGCCATAGATAAAATAAAAGATTTTATTAGTCTCCAGAAAAGGGGGGAATGAAG 1685
QY 3928 ACCCACCTGAGTGTGGCAAGC-----TAGCTTAAAGTAAGCCCAT 3969
Db 1686 ACCCACCTGAGTGTGGCAAGCCTAGAGTCTGATAGCCCGCATAGCCCAT 1745
QY 3970 TTTCAAGGCGATGG-AAAATAATACTGAGATAGAGAGTTTCAAGTCAAGGTTAGGAA 4028
Db 1746 TTTCAAGGCGATGG-AAAATAATACTGAGATAGAGAGTTTCAAGTCAAGGCGGTA 1805
QY 4029 CAGAGACACAGAAATATGGGCCAAACAGAGATATCTGTGTAGAGATTTCTGCCCCGC 4088
Db 1806 CATGAAAATAGCTAACGTTGGGCCAAACAGAGATATCTGCGGTGAGCAGTTTCGCGCCGG 1865
QY 4089 TCAGGGCCAGAACAGATTGGAAACAGGAGAAATATGGGCCAAACAGAGATATCTGTGTAGC 4148
Db 1866 CCGGGGCAAGAACAGATGG-----TCACCGC 1892
QY 4149 AGTTCTCTCCCGCGCTCAGGGCCCAAGAACAGATGTTCCCGAGATCGGTCGCCGCTCAG 4208
Db 1893 AGTTCTCGCCCGCGCGGAGGCAAGAACAGATGTTCCCGAGATATGCCCCAACCTCAG 1952
QY 4209 CAGTTCTTAGAAGAACCATCAGATGTTCCAGGGTGGCCCAAGGACCTGAATGACCTGT 4268
Db 1953 CAGTTCTTAGAAGAACCATCAGATGTTCCAGGGTGGCCCAAGGACCTGAATGACCTGT 2012
QY 4269 GCCTTATTGAACTAAACCAATCAGTTCGCTTCTGCTTCTGTCGCGGCTTCTGCTCCC 4328
Db 2013 GCCTTATTGAACTAAACCAATCAGTTCGCTTCTGCTTCTGTCGCGGCTTCTGCTCCC 2072
QY 4329 CGAGCTCAATAAAGAGCCCAACACCTCACTCGGGCGCGCAGTCTCTCCGATAGACTGC 4388
Db 2073 CGAGCTCTATAAAGAGCTCACACACCTCACTCGGGCGCGCAGTCTCTCCGATGACTGA 2132
QY 4389 GTCCCGCGGTACCCGTTGTTCTCAATAAAACCTCTTGCAGTTGCATCCGATCGTGTCT 4448

```



Db 2133 GTCCCGGGTACCCGGTTCCTCAATAAACCCCTCTTGCACTGTCATCCGACTCGTGGTCT 2192  
Qy 4449 CGCTGTTCTTGGGGGGTCT-CTCTGAGTGAATGACTACCCGTGAGC-GGGGTCTTTTCA 4506  
Db 2193 CGCTGTTCTTGGGGGGTCTCTCTGAGTGAATGACTACCCGTGAGCGGGGTCTTTTCA 2252  
Qy 4507 GTTCTTCCCACTACACAGTCTCACTAAACATTCCTGATGTGCGGAGGAGCTCCGTGAG 4566  
Db 2253 GTTCTTCCCACTACACAGTCTCACTAAACATTCCTGATGTGCGGAGGAGCTCCGTGAG 2312  
Qy 4567 CCGGGTTTTGTTTAAATAAATGCAAGAACAGTGTTCCTTTCAAGCCAGACTACATCC 4626  
Db 2313 CCGGGTTTGTGTTTAAATAAATGCAAGAACAGTGTTCCTTTCAAGCCAGACTACATCC 2372  
Qy 4627 TGAATCTCGGCTTTAAAGAGATGTGAAGGCTCTGTGAGTATCTGGAATCTGCGCACAGACTT 4686  
Db 2373 TGAATCTCGGCTTTAAAGAGATGTGAAGGCTCTGTGAGTATCTGGAATCTGCGCACAGACTT 2431  
Qy 4687 TTTAAGATTTTATGCTCTCTGATGAGGATTTAGTCAATCTATCTCTGCTCTATTTTGC 4746  
Db 2432 TTTAAGATTTTATGCTCTCTGATGAGGATTTAGTCAATCTATCTCTGCTCTATTTTGC 2491  
Qy 4747 TGGCTTCTCGTATTTTAAATTTCTAGTTTGCACTCCCTTCTGAGAGCACGGCATTCG 4806  
Db 2492 TGGCTTCTCGTATTTTAAATTTCTAGTTTGCACTCCCTTCTGAGAGCACGGCATTCG 2551  
Qy 4807 AGAGTAGTTAATCTCTGAGGCGAGGCTTCTGTGAAGAGTTGCTTGGGCTCAGTGTGAG 4866  
Db 2552 AGAGTAGTTAATCTCTGAGGCGAGGCTTCTGTGAAGAGTTGCTTGGGCTCAGTGTGAG 2611  
Qy 4867 ATTTTCCCAATAAAGGGGCTTGCCTCTGTGTACAGACAGATCCGAATCTAGAGTGCAT 4926  
Db 2612 ATTTTCCCAATAAAGGGGCTTGCCTCTGTGTGTACAGACAGATCCGAATCTAGAGTGCAT 2671  
Qy 4927 ACTCAGAGTCCCGGGTTCCGGGGCTGTGATCTCAGGGCATCTTTGGCTAGAGATCCCTC 4986  
Db 2672 ACTCAGAGTCCCGGGTTCCGGGGCTGTGATCTCAGGGCATCTTTGGCTAGAGATCCCTC 2731  
Qy 4987 TAGCCGGAGCATCTGTCGGCGGTACCGAGCTCGAATTCGTAATCATGGTCAATAGCTGT 5046  
Db 2732 TAGCCGGAGCATCTGTCGGCGGTACCGAGCTCGAATTCGTAATCATGGTCAATAGCTGT 2791  
Qy 5047 TTC-----CTGTGTGAATTTGTATTCGCTCACAATTCACACACATACGA 5093  
Db 2792 ATGCCGACATCACCCATGGGAGATCGGGTCCGCACTTCGGGCTCATGAGCGGTTGT 2851  
Qy 5094 GCCGAAGCATAAAGTGTAAAGCTCGGGGTGCTTAATGAGT----- 5134  
Db 2852 TTCGGGTGGTATGTCGAGGCCCCGGTGGCGGGGACTGTGGGGCCCATCTCCTTG 2911  
Qy 5135 -----GAGCTAACTACATTAATTTGCGTTCGCTCACTGCC 5170  
Db 2912 CATGACCATTCCTTGGCGGGCGGTGCTCAAGCGGCTCAACCTACTAGGGCTGCTTC 2971  
Qy 5171 CGCTTTCAGTCGGGAAACCTGTGTCAGCTGCAATTAATGAATCGGCAACCGCGGG 5230  
Db 2972 CTAATCAGAGTCGCATTAAGGAGAGCGTCTGCAATTAATGAATCGGCAACCGCGGG 3031  
Qy 5231 GAGAGCGGTTTGGGCGCTTTCCGCTTCTCGCTCACTGACTCGCTCGGCTC 5290  
Db 3032 GAGAGCGGTTTGGGCGCTTTCCGCTTCTCGCTCACTGACTCGCTCGGCTC 3091  
Qy 5291 GGTCTGTTCCGTCGGCGAGCGGTATCAGCTCACTCAAAAGGCGGTAAATCGGTATCCAC 5350  
Db 3092 GGTCTGTTCCGTCGGCGAGCGGTATCAGCTCACTCAAAAGGCGGTAAATCGGTATCCAC 3151  
Qy 5351 AGAATCAGGGATTAACGAGGAAAGAACATGTGAGCAAAAGGCGCAAAAGCCAGGAA 5410  
Db 3152 AGAATCAGGGATTAACGAGGAAAGAACATGTGAGCAAAAGGCGCAAAAGCCAGGAA 3211  
Qy 5411 CCGTAAAAAGGCGGCTTGTGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCA 5470  
Db 3212 CCGTAAAAAGGCGGCTTGTGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCA 3271

Qy 5471 CAAAATCGACGCTCAAGTCAAGGTGGGAAACCGGACGACTATAAAGTACACAGC 5530  
Db 3272 CAAAATCGACGCTCAAGTCAAGGTGGGAAACCGGACGACTATAAAGTACACAGC 3331  
Qy 5531 GTTTCCTCCCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCCCTGCCGCTTACCGGATA 5590  
Db 3332 GTTTCCTCCCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCCCTGCCGCTTACCGGATA 3391  
Qy 5591 CTTGTCGCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGATAGTA 5650  
Db 3392 CTTGTCGCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGATAGTA 3451  
Qy 5651 TCTCAGTTCCGGTGTAGTTCGCTCCAAAGCTGGGCTGTGTGACGACCCCGCTTCA 5710  
Db 3452 TCTCAGTTCCGGTGTAGTTCGCTCCAAAGCTGGGCTGTGTGACGACCCCGCTTCA 3511  
Qy 5711 GCCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAAACCCCGGTAAAGACGA 5770  
Db 3512 GCCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAAACCCCGGTAAAGACGA 3571  
Qy 5771 CTTATCGGCACCTGGCAGCAGCCACTGGTAAACAGGATTTAGCAGCAGGATGTAGGCGG 5830  
Db 3572 CTTATCGGCACCTGGCAGCAGCCACTGGTAAACAGGATTTAGCAGCAGGATGTAGGCGG 3631  
Qy 5831 TGCCTACAGATTCTTGAAGTGGTGGCTTAACTACCGCTACACTAGAGGACAGTATTTGG 5890  
Db 3632 TGCCTACAGATTCTTGAAGTGGTGGCTTAACTACCGCTACACTAGAGGACAGTATTTGG 3691  
Qy 5891 TATCTGCGCTCTGCTGAAAGCCAGTTACCTTCGGAAGGAGTTGGTAGCTCTTGATCCCG 5950  
Db 3692 TATCTGCGCTCTGCTGAAAGCCAGTTACCTTCGGAAGGAGTTGGTAGCTCTTGATCCCG 3751  
Qy 5951 CAAAACAAACCCGCTGTGAGCGTGGTGTGTTTGTTCGAAAGCAGAGATTAACGCGAG 6010  
Db 3752 CAAAACAAACCCGCTGTGAGCGTGGTGTGTTTGTTCGAAAGCAGAGATTAACGCGAG 3811  
Qy 6011 AAAAAAGGATCTCAAGAGATCCCTTGTGATCTTTTCTAGCGGCTGACGCTCAGTGGA 6070  
Db 3812 AAAAAAGGATCTCAAGAGATCCCTTGTGATCTTTTCTAGCGGCTGACGCTCAGTGGA 3871  
Qy 6071 CGAAACTCAGCTTAAAGGATTTTGTCTATGAGATTAACAAGGATCTTCAACCTAGAT 6130  
Db 3872 CGAAACTCAGCTTAAAGGATTTTGTCTATGAGATTAACAAGGATCTTCAACCTAGAT 3931  
Qy 6131 CTTTAAATTTAAATGAAATTTTAAATCAATCTAAAGTATATATGATGATAAATTTGTC 6190  
Db 3932 CTTTAAATTTAAATGAAATTTTAAATCAATCTAAAGTATATATGATGATAAATTTGTC 3991  
Qy 6191 TGAAGTTACCAATGCTTAATCAGTGAGGCACTTCTCAGCGATCTGCTATTTGTTTC 6250  
Db 3992 TGAAGTTACCAATGCTTAATCAGTGAGGCACTTCTCAGCGATCTGCTATTTGTTTC 4051  
Qy 6251 ATCCATAGTTGCTGACTCCCCGCTGCTAGATAACTACGATACGGGAGGCTTACCATC 6310  
Db 4052 ATCCATAGTTGCTGACTCCCCGCTGCTAGATAACTACGATACGGGAGGCTTACCATC 4111  
Qy 6311 TGGCCCCAGTGTGCAATGATACCCGAGACCCACGCTCACCGGCTCCAGATTTATFAGC 6370  
Db 4112 TGGCCCCAGTGTGCAATGATACCCGAGACCCACGCTCACCGGCTCCAGATTTATFAGC 4171  
Qy 6371 AATAAACAGCCAGCGGAGGCGGCGAGAGAGTGGTCTGCAATTTATTCGCTTC 6430  
Db 4172 AATAAACAGCCAGCGGAGGCGGCGAGAGAGTGGTCTGCAATTTATTCGCTTC 4231  
Qy 6431 CATCCAGTCTATTAATTTGTCGGGAAAGCTAGAGTAAAGTGTTCGCCAGTTAATAGTTT 6490  
Db 4232 CATCCAGTCTATTAATTTGTCGGGAAAGCTAGAGTAAAGTGTTCGCCAGTTAATAGTTT 4291  
Qy 6491 GCGCAACGTTGTGCAATGCTTACAGGC-TGCTGTGTACGCTCGTGTGTTGTTGATGCG 6549  
Db 4292 GCGCAACGTTGTGCAATGCTTACAGGCATGCTGTGTGTCAGCTCGTGTGTTGATGCG 4351



1982 CAGTTTCTTAAAGACCCATCAGATGTTTCCAGGCTCCCCAAGGACCTGAAATGACCCCTGC 2041  
 4269 GCCTTATTTGAATTAACCAATCAGTTTCGCTTCTCGCTTCTGTTTGGCGCTTCTGCTCC 4328  
 2042 GCCTTATTTGAATTAACCAATCAGCTTCTTCTGCTTCTGTTTGGCGCTTCTGCTCC 2101  
 4329 CGAGCTCAATAAAGAGCCCAACACCCCTCACTCGCGCGCCAGTCCCTCCGATAGACTGC 4388  
 2102 CGAGCTCATAAAGAGCTCAACACCCCTCACTCGCGCGCCAGTCCCTCCGATGACTGA 2161  
 4389 GTCCGCGGGTACCGGTGTTCTCAATAAACCCTTTGCAAGTTGCAATCCGACTCGTGGTCT 4448  
 2162 GTCCGCGGGTACCGGTGTTCTCAATAAACCCTTTGCAAGTTGCAATCCGACTCGTGGTCT 2221  
 4449 CGCTGTTCTTGGGAGGCTCTCTCTGAGTGATGACTACCCGTCAGC-GGGGCTTTTCA 4506  
 2222 CGCTGTTCTTGGGAGGCTCTCTCTGAGTGATGACTACCCGTCAGCGGGGCTTTTCA 2281  
 4507 GTTCTTCCCACTACACAGGTCTCACTAAATCTCTGATGTGCGCAGGGAATCCGTCAG 4566  
 2282 GTTCTTCCCACTACACAGGTCTCACTAAATCTCTGATGTGCGCAGGGAATCCGTCAG 2341  
 4567 CCGGTTTTTGTATTAATAAATGCAAGAACAGTGTTCCTTTCAAGCCAGACTACATCC 4626  
 2342 CCGGTTTTTGTATTAATAAATGCAAGAACAGTGTTCCTTTCAAGCCAGACTACATCC 2401  
 4627 TCACTCTCGGCTTATAAAGATGTTGAAGGCTCTGGAATCTATCTGCACTATCTGCAAGACTT 4686  
 2402 TCACTCTCGGCTTATAAAGATGTTGAAGGCTCTGGAATCTATCTGCACTATCTGCAAGACTT 2460  
 4687 TTTAAGATTTTATGCTCTCGATGAGGATTTAGTCAATCTATCTCTGCTATTTTGC 4746  
 2461 TTTAAGATTTTATGCTCTCGATGAGGATTTAGTCAATCTATCTCTGCTATTTTGC 2520  
 4747 TGGCTTCTCCGATTTTAAATTTCTAGTTTGCATCTCCCTTCTGAGAGCAGCGGATTC 4806  
 2521 TGGCTTCTCCGATTTTAAATTTCTAGTTTGCATCTCCCTTCTGAGAGCAGCGGATTC 2580  
 4807 AGAGTAGTTTAACTCTGAGGAGGCTCTGTAAGAGGTTGCTGGGCTCAGTGTGAG 4866  
 2581 AGAGTAGTTTAACTCTGAGGAGGCTCTGTAAGAGGTTGCTGGGCTCAGTGTGAG 2640  
 4867 ATTTTGGCCATAAAGGGGCTCTGCCCCCTGTGTACAGACAGATCGGAATCTAGAGTGCAT 4926  
 2641 ATTTTGGCCATAAAGGGGCTCTGCCCCCTGTGTACAGACAGATCGGAATCTAGAGTGCAT 2700  
 4927 ACTCAGAGTCCCGGGTTCGCGGCTCTGATCTCAGGGCATCTTTGCTCAGAGATTCCTC 4986  
 2701 ACTCAGAGTCCCGGGTTCGCGGCTCTGATCTCAGGGCATCTTTGCTCAGAGATTCCTC 2760  
 4987 TAGCCGGAACGATCGTGGCGGGTACCGAGCTCGAATTCGTAATCATGCTCATAGCTGT 5046  
 2761 TAGCCGGAACGATCGTGGCGGGTACCGAGCTCGAATTCGTAATCATGCTCATAGCTGT 2820  
 5047 TTTCTCTGTTGAAATTTCTGTTATCTGCTTCAATTTCCGTCACAAATTCACACAACTACGA 5093  
 2821 ATCCGCGACATCACCGATGGGGAAGATCGGGTTCGCCACTTCGGGCTCATGAGCGCTGT 2880  
 5094 GCGGGAAGCATAAAGTGTAAAGCTGGGGTGGCTTAATGAGT----- 5134  
 2881 TTTGGCGTGGTATGGTGGCAGGCCCGGTGGCGGGGACTGTGGGGCCATCTCCTTTG 2940  
 5135 -----GAGCTAACTCAATTAATTTGGCTTGGCTCAGTGC 5170  
 2941 CATGACCAATTCCTTGGCGGGGCTGCTCAAGCGGCTCAACCTACTACTGGGCTGCTTC 3000  
 5171 CGCTTTCCAGTCGGGAACCTGTGCGCCAGCTGCAATTAATGAATCGGCAACGCGCGG 5230  
 3001 CTAATCAGAGTCCGATTAAGGAGAGCGCTCTGCAATTAATGAATCGGCAACGCGCGG 3060  
 5231 GAGAGCGGTTTGGTATTTGGGCGCTTTCCGCTTCTCGCTCACTGACTCGCTGCGCTC 5290  
 3061 GAGAGCGGTTTGGTATTTGGGCGCTTTCCGCTTCTCGCTCACTGACTCGCTGCGCTC 3120

QY 5291 GGTCTTGGCTCGCGGAGCGGTATCAGCTCACTCAAAGCGGTAAATACCGTTATCCAC 5350  
 DB 3121 GGTCTTGGCTCGCGGAGCGGTATCAGCTCACTCAAAGCGGTAAATACCGTTATCCAC 3180  
 QY 5351 AGAATCAGGGATTAACGCAAGAAACATGTGAGCAAAAGCCAGCAAAAGCCAGGAA 5410  
 DB 3181 AGAATCAGGGATTAACGCAAGAAACATGTGAGCAAAAGCCAGCAAAAGCCAGGAA 3240  
 QY 5411 CGGTAAAGGCGGCTGCTGGGCTTTTCCATAGGCTCGCCCCCTGAGAGCATCA 5470  
 DB 3241 CGGTAAAGGCGGCTGCTGGGCTTTTCCATAGGCTCGCCCCCTGAGAGCATCA 3300  
 QY 5471 CAAATAACGACGCTCAAGTCAGAGGTGCGGAAACCCGACAGGACTATAAAGATACAGGC 5530  
 DB 3301 CAAATAACGACGCTCAAGTCAGAGGTGCGGAAACCCGACAGGACTATAAAGATACAGGC 3360  
 QY 5531 GTTTCCTCCCTGGAAAGCTCCCTCGTGCCTCTCTGTTCCGACCTTCCGCTTACCGGATA 5590  
 DB 3361 GTTTCCTCCCTGGAAAGCTCCCTCGTGCCTCTCTGTTCCGACCTTCCGCTTACCGGATA 3420  
 QY 5591 CTTGTCGCGCTTTCTCCCTTCCGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGTA 5650  
 DB 3421 CTTGTCGCGCTTTCTCCCTTCCGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGTA 3480  
 QY 5651 TCTCAGTTCGCTGTAGTCTGCTCCAAAGCTGGGCTGTGTGACAGCAACCCCGCTTCA 5710  
 DB 3481 TCTCAGTTCGCTGTAGTCTGCTCCAAAGCTGGGCTGTGTGACAGCAACCCCGCTTCA 3540  
 QY 5711 GCCGACCGCTGCGCTTATCCGTAACCTATCGTCTCAGTCCAAACCCCGGTAAGACAGA 5770  
 DB 3541 GCCGACCGCTGCGCTTATCCGTAACCTATCGTCTCAGTCCAAACCCCGGTAAGACAGA 3600  
 QY 5771 CTTATCGGCACTGGCAGCAGCACTGGTAAACAGGATTAAGCAGCGAGTATGAGGCGG 5830  
 DB 3601 CTTATCGGCACTGGCAGCAGCACTGGTAAACAGGATTAAGCAGCGAGTATGAGGCGG 3660  
 QY 5831 TGCTACAGATTTCTGAAAGTGTGGCTTAACTTACGGCTACACTAGAAGACAGATTTTGG 5890  
 DB 3661 TGCTACAGATTTCTGAAAGTGTGGCTTAACTTACGGCTACACTAGAAGACAGATTTTGG 3720  
 QY 5891 TATCTGCGCTCTGCTGAAAGCAGTTTACCTTCGGAAGGAGTTGGTGTCTTGTATCCGG 5950  
 DB 3721 TATCTGCGCTCTGCTGAAAGCAGTTTACCTTCGGAAGGAGTTGGTGTCTTGTATCCGG 3780  
 QY 5951 CAAACAAACCAACCGCTGTAGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 6010  
 DB 3781 CAAACAAACCAACCGCTGTAGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3840  
 QY 6011 AAAAAAGGATCTCAAGAGATCTTTCATCTTTTCTACGGGCTCTGACGCTCAGTGGAA 6070  
 DB 3841 AAAAAAGGATCTCAAGAGATCTTTCATCTTTTCTACGGGCTCTGACGCTCAGTGGAA 3900  
 QY 6071 CGAAACTCAGCTTAAAGGATTTTGGTCAATGAGATTAATAAGGATTTTCAACCTAGAT 6130  
 DB 3901 CGAAACTCAGCTTAAAGGATTTTGGTCAATGAGATTAATAAGGATTTTCAACCTAGAT 3960  
 QY 6131 CCTTTTAAATTAATAAGGATTTTAAATCAATCTAAAGATATATAGATTAACCTTGGTC 6190  
 DB 3961 CCTTTTAAATTAATAAGGATTTTAAATCAATCTAAAGATATATAGATTAACCTTGGTC 4020  
 QY 6191 TGACAGTACCAATGCTTAAACAGTGAGGCACTTATCTCAGCGATCTGTCTATTTCTGTC 6250  
 DB 4021 TGACAGTACCAATGCTTAAACAGTGAGGCACTTATCTCAGCGATCTGTCTATTTCTGTC 4080  
 QY 6251 ATCCATAGTGTGCTGACTCCCGCTGTGTAGATAACTACGATACGGGAGGCTTACCATC 6310  
 DB 4081 ATCCATAGTGTGCTGACTCCCGCTGTGTAGATAACTACGATACGGGAGGCTTACCATC 4140  
 QY 6311 TGGCCCCAGTGTGCAATGATACCGGAGACCCAGCTTCCAGGCTCCAGATTTATCAGC 6370  
 DB 4141 TGGCCCCAGTGTGCAATGATACCGGAGACCCAGCTTCCAGGCTCCAGATTTATCAGC 4200

QY 6371 AATAACAGCCAGCGGAGGCGGAGAGTGGTCTGCAACTTTATCGGCTC 6430  
Db 4201 AATAACAGCCAGCGGAGGCGGAGAGTGGTCTGCAACTTTATCGGCTC 4260  
QY 6431 CATCAGTCTAATAATTGTCGCGGAGCTAGAGTAAGTTCGCCAGTTAATAGTTT 6490  
Db 4261 CATCAGTCTAATAATTGTCGCGGAGCTAGAGTAAGTTCGCCAGTTAATAGTTT 4320  
QY 6491 GCGCAACGTTGTGCAATGCTACAGGC-TGGTGGTGTACGCTCGTGGTTGGTATGGC 6549  
Db 4321 GCGCAACGTTGTGCAATGCTACAGGCATCGTGGTGTACGCTCGTGGTTGGTATGGC 4380  
QY 6550 TTCAATCAGCTCCGGTTCGCCAAGATCAAGGGAGTTACATGATCCCCCATGTTGTGCA 6609  
Db 4381 TTCAATCAGCTCCGGTTCGCCAAGATCAAGGGAGTTACATGATCCCCCATGTTGTGCA 4440  
QY 6610 AAAAGCGGTAGCTCTTCGGTCTCCGATCGTGTGCAAGTAAGTTGGCGCGAGTGT 6669  
Db 4441 AAAAGCGGTAGCTCTTCGGTCTCCGATCGTGTGCAAGTAAGTTGGCGCGAGTGT 4500  
QY 6670 ATCACTCATGTTATGGCAGCACTGCATAATTTCTTACTGTCTATGCCATCCGTAAGATG 6729  
Db 4501 ATCACTCATGTTATGGCAGCACTGCATAATTTCTTACTGTCTATGCCATCCGTAAGATG 4560  
QY 6730 CTTTTCTGTGACTGTGAGTACTCAACCAAGTCAATCTTGAGAAATGATGTGGCGGAC 6789  
Db 4561 CTTTTCTGTGACTGTGAGTACTCAACCAAGTCAATCTTGAGAAATGATGTGGCGGAC 4620  
QY 6790 GAGTGTCTTTCCCGGCGTCAATACGGGTAATACCGGCGCACATAGCAACTTTAA 6849  
Db 4621 GAGTGTCTTTCCCGGCGTCAATACGGGTAATACCGGCGCACATAGCAACTTTAA 4680  
QY 6850 AGTGTCTCATCTTGAAACGTTCTTCGGGCGGAAACTCTCAAGGATCTTACCGCTGT 6909  
Db 4681 AGTGTCTCATCTTGAAACGTTCTTCGGGCGGAAACTCTCAAGGATCTTACCGCTGT 4740  
QY 6910 GAGATCCAGTTTCGATGTAACCACTCGTGCACCCCACTGATCTTACGATCTTTACTTT 6969  
Db 4741 GAGATCCAGTTTCGATGTAACCACTCGTGCACCCCACTGATCTTACGATCTTTACTTT 4800  
QY 6970 CACGAGCTTTCTGGGTGAGCAAAAAGGAGGCAAAATCGCGCAAAAAGGGAATAG 7029  
Db 4801 CACGAGCTTTCTGGGTGAGCAAAAAGGAGGCAAAATCGCGCAAAAAGGGAATAG 4860  
QY 7030 GCGCAGCGGAAATCTTCAATCTCATCTCTCTCTTTTCAATATTTGAAGCATTTA 7089  
Db 4861 GCGCAGCGGAAATCTTCAATCTCATCTCTCTCTTTTCAATATTTGAAGCATTTA 4920  
QY 7090 TCAGGGTTATTTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACCAAT 7149  
Db 4921 TCAGGGTTATTTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACCAAT 4980  
QY 7150 AGGGGTTCCGGGCACATTTCCCGGAAAGTGCACCTGACGTCTAAGAAACCATTTAT 7209  
Db 4981 AGGGGTTCCGGGCACATTTCCCGGAAAGTGCACCTGACGTCTAAGAAACCATTTAT 5040  
QY 7210 CATGACATTAACCTATAAATAGCGGTATCAGAGGCGCTTCTGCTCGCGCGTTTCGG 7269  
Db 5041 CATGACATTAACCTATAAATAGCGGTATCAGAGGCGCTTCTGCTCGCGCGTTTCGG 5100  
QY 7270 TGATGACCGGTGAAACCTCTGACACATGACGCTCCCGGAGACGGTCAAGCTTTGTCTGA 7329  
Db 5101 TGATGACCGGTGAAACCTCTGACACATGACGCTCCCGGAGACGGTCAAGCTTTGTCTGA 5160  
QY 7330 AGCGGATCCGGGAGCAGCAAGCCGTCAGGGCGCGTTCAGCGGTTGTTGGCGGTGTCG 7389  
Db 5161 AGCGGATCCGGGAGCAGCAAGCCGTCAGGGCGCGTTCAGCGGTTGTTGGCGGTGTCG 5220  
QY 7390 GGGCTGCTTAATCTATGGGATCAGACAGATTTACTGAGTGTGACCATATTCGGTG 7449  
Db 5221 GGGCTGCTTAATCTATGGGATCAGACAGATTTACTGAGTGTGACCATATTCGGTG 5280  
QY 7450 TGAAATACCGCAGATCGTAAAGGAGAAATACCGCATCAGG 7492

Db 5281 TGAAATACCGCAGATCGTAAAGGAGAAATACCGCATCAGG 5323  
RESULT 6  
US-08-336-132-1  
; Sequence 1, Application US/08336132  
; Patent No. 5693508  
; GENERAL INFORMATION:  
; APPLICANT: CHANG, LUNG-JI  
; TITLE OF INVENTION: RETROVIRAL VECTORS  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HAVERSTOCK, MEDLEN & CARROLL  
; STREET: 220 MONTGOMERY STREET, SUITE 2200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/336,132  
; FILING DATE: 07-NOV-1994  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CARROLL, PETER G.  
; REGISTRATION NUMBER: 32,837  
; REFERENCE/DOCKET NUMBER: CHANG-00817  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6145 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-336-132-1  
Query Match 40.0%; Score 3062.2; DB 1; Length 6145;  
Best Local Similarity 71.7%; Pred. No. 0;  
Matches 4818; Conservative 0; Mismatches 1168; Indels 735; Gaps 32;  
QY 549 TTTTGAAGACCCCACTGTAGGTTTGGCAAGCTAGCTTAAAGTAAAGCCATTTTGAAGG 608  
Db 144 TTTTGAAGACCCCACTGTAGG--TGGCAAGCTAGCTTAAAGTAAAGCCACTTTGCAAGG 201  
QY 609 CATGG-AAAATACATACTGAGATAGAGAAGTTTCAGATCAAGTTAGGAACAGAGAC 667  
Db 202 CATGGAATAATACATACTGAGATAGAGAAGTTTCAGATCAAGTTAGGAACAGAGAAC 261  
QY 668 AGCAAGATATGGGCGCAACAGAGATATCTGTGTAGAGATTTCTTGGCCCC-GCTCAGGGCC 726  
Db 262 AGCTGAATA-----CCAAACAGGATATCTGTGTAGCGGTTCTTGGCCCCGCTCAGGGCC 317  
QY 727 AAGAACAGTTTGGAAACAGGAGAAAT-ATGGGCGCAACAGGAGATATCTGTGTAGAGAGTTTCT 785  
Db 318 AAGAACAGATGAGACAGCTGATGATGGCCCAACAGGATATCTGTGTAGAGAGTTTCT 377  
QY 786 GCGCCGGCTCAGGCGCCAGAAACAGATGTTCCCGAGATGCGGTCCCGCCCTCAGCAGTTTC 845  
Db 378 GCGCCGGCTCAGGCGCCAGAAACAGATGTTCCCGAGATGCGGTCCCGCCCTCAGCAGTTTC 437  
QY 846 TAGAGAACCATCAGATGTTTCCAGGGTGCCTCCCAAGGACCTG-AAATGACCTGTGCTTTA 904  
Db 438 TAGTGAATCATCAGATGTTTCCAGGGTGCCTCCCAAGGACCTGAAAATGACCTGTACCTTTA 497  
QY 905 TTTGAACCTAACCAATCAGTTCTGCTTCTGTTTCTGCGCGCTTCTGCTCCCGGAGCT 964





QY 3122 CCGGGTCTATTTTGTGCAAGCCTTACTTCGGCTTCCCCTGGTTTGTCTTATTTGGGGCC 3181  
DB |||||  
2562 GGGACGCC-----GGCTGGATGATCCT 2583  
QY 3182 AAGGACCCCGGTCAACCGTCTCCAGTGTCAAGCCACACGACGCGCGCGACAC 3241  
DB |||||  
2584 CCAGCGCGGGATCTCATCTGGAGTTCTTCGCCACCCCGGGCTCGATCCCTCGCGAG 2643  
QY 3242 CAACACCGCGGCCCAATCGCGTCCGACGCCCTGTCCCTGCGCCACAGAGCGGCTCGGC 3301  
DB |||||  
2644 TTGGTTACGTCTGCTGAGGCTGGACGACCTCG----- 2678  
QY 3302 CAGCGCGCGGGCGCGAGTGCACACGAGGGGCTGGACTTCGCCCTCGATGCCAACTCT 3361  
DB |||||  
2679 -----CGAGTTCTACCGCGCAGTGCATAATCCG 2705  
QY 3362 GCTACCTCTGGATGGAATCTCTTCACTATGATGTCACTTCTCACTGCCTTGTCTCTGA 3421  
DB |||||  
2706 TCGGATCCAGCA---AACACGACGGCTATCCCGGCATCATGCCCCCGAACTGCAGG 2762  
QY 3422 GAGTGAAGTTACGACGAGCGCAGAGCCCGCGGTACGACGAGGGCCAGAACACGACTCT 3481  
DB |||||  
2763 AGTGGGAGGCAGATGCGCGCTTTGGTTCGACCCCGACGGGACGCTCTGCGCCTGTATAC 2822  
QY 3482 ATAACGAGTCAATCTAGGAAGAAGAGAGGTACGATGTTTGGACAAGAGACGTGCC 3541  
DB |||||  
2823 AGAACGAATGCTTCAGCGCATCTCATGAGTGTCTT-----CCCGTTTTC 2869  
QY 3542 GGGACCTGAGATGGGGGGAAGCCGAGAGGAAGAACCTTCAGGAAGCCCTGTACAAATG 3601  
DB |||||  
2870 CGCTGAGTCACTCGCGTGGATGGAGCGCTGCGCCTCTCGCGACCGCGAG---CTG 2925  
QY 3602 AACTGCAGAAAGATAAGATGGCGAGGCTTACAGTGAGATTGGGATGAAGCGCGAGCGCC 3661  
DB |||||  
2926 CTCACCACTCGAGGGGCTGACGCTGCAGAGCGCGAGTCGAACTGCTG----- 2978  
QY 3662 GGAGGGCAAGGGGCAAGATGGCTTTTACAGGCTCTCAGTACAGCCACCAAGGACACT 3721  
DB |||||  
2979 -----TCCAAAGGGACCTCAAGGCTTTCGAGGGACACTAGGCTGACTCCATCGAGCCAG 3033  
QY 3722 ACAGCGCCCTTACATGACGAGCCCTGCCCCCTCGCTAACTCGACGCGCGCGGATCCGG 3781  
DB |||||  
3034 TGTA-----GAGATAAGCTTATCG 3052  
QY 3782 ATTAGTCCAAATTTGTTAAAGACAGATATCAGTGTCCAGGCTCTAGTTTGTACTCAACA 3841  
DB |||||  
3053 ATTAGTCCAAATTTGTTAAAGACAGATATCAGTGTCCAGGCTCTAGTTTGTACTCAACA 3112  
QY 3842 ATATCACAGCTGAAGCTTATAGATACGAGCCATAGATAAAATAAAAGATTTTATTAG 3901  
DB |||||  
3113 ATATCACAGCTGAAGCTTATAGATACGAGCCATAGATAAAATAAAAGATTTTATTAG 3172  
QY 3902 TCTCAGAAAAGGGGGAATGAAGACCCACCTGTAGTTGGTGGCAAGCTAGCTTAAGT 3961  
DB |||||  
3173 TCTCAGAAAAGGGGGAATGAAGACCCACCTGTAGTTGGTGGCAAGCTAGCTTAAGT 3232  
QY 3962 AACGCCATTTTGAAGGATGG-AAAATACATAACTGAGATAGAGAGTTTCAGATCAAG 4020  
DB |||||  
3233 AACGCCATTTTGAAGGATGGAAAATAACATACTGAGATAGAGAGTTTCAGATCAAG 3292  
QY 4021 GTTAGGAACAGA-GAGACAGCAGATATGGGCCAAACAGGATATCTGTGGTAAGCAGTTTC 4079  
DB |||||  
3293 GTCAGGAACAGATGGAACAGCTGAATAAGGCCAAACAGGATATCTGTGGTAAGCAGTTTC 3352  
QY 4080 CTGCCCC--GCTCAGGGCCAAAGACGTTGGAAACAGGAGATATGGGCCAAACAGGATATC 4138  
DB |||||  
3353 CTGCCCCGGCTCAGGGCCAAAGACAGATGGAAACAGCTGAATATGGGCCAAACAGGATATC 3412  
QY 4139 TGTGTAGACGATTTCTCCCGGCTCAGGGCCAAAGACAGATGTTCCCGAGTCCGCTC 4198  
DB |||||  
3413 TGTGTAGACGATTTCTCCCGGCTCAGGGCCAAAGACAGATGTTCCCGAGTCCCGCTC 3472

QY 4199 CCGCCCTCAGACGATTTCTTAGAGAACCATCAGATGTTTTCAGGGTGCCTCCCAAGACCTGAA 4258  
DB |||||  
3473 CAGCCCTCAGCAGATTTCTTAGAGAACCATCAGATGTTTTCAGGGTGCCTCCCAAGACCTGAA 3532  
QY 4259 ATGACCCCTGCTCTTATTTGAACATAACAGTTCCGTTCTCGCTTCTGTTCCGGCGC 4318  
DB |||||  
3533 ATGACCCCTGCTCTTATTTGAACATAACAGTTCCGTTCTCGCTTCTGTTCCGGCGC 3592  
QY 4319 TTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGCGCGCGCAGCTCTCC 4378  
DB |||||  
3593 TTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGCGCGCGCAGCTCTCC 3652  
QY 4379 GATAGACTGCGTCCCGCGGTACCGTGTCTCAATAAACCCTCTTGACAGTTGCATCCGA 4438  
DB |||||  
3653 GATTGACTGAGTCCCGCGGTACCGTGTATCAATAAACCCCTTTGACAGTTGCATCCGA 3712  
QY 4439 CTGCTGCTCTCGCTCTTCTTGGGAGGCTCT-CTCTGAGTGTATGACTACCGCTCAGCG 4497  
DB |||||  
3713 CTGCTGCTCTCGCTCTTCTTGGGAGGCTCTCTCTGAGTGTATGACTACCGCTCAGCG 3772  
QY 4498 GGTCTTTTCTGCTCTTCTCCACCTTACACAGTCTCACTAAACATTTCTGATGTGCCGAGGA 4557  
DB |||||  
3773 GGTCTTTTCAATTTGGGGCTCGTCCGGATCGGAGACCCCTGCCCGGAGCACCGACC 3832  
QY 4558 CTCCGTACGCGCGTCTTGTATTAATAAATGAAGAACAGTGTCTTCAAGCCAG 4617  
DB |||||  
3833 CACCACCGGAGGTAAGCTG----- 3852  
QY 4618 ACTACATCTGACTCTCGGCTTTATAAAGATGTTGAAGGCTCTGTGGACTATCTGCC 4677  
DB |||||  
3853 -----GCTCCCTCGCGCTTTCGGTATGACCGTGAACCCCTCTGAC-ACATCAGCT 3904  
QY 4678 ACAGACTTTTAAAGATTTTATGCTCTCGATGAGGATTTAGTCAATCTATCTCTCT 4737  
DB |||||  
3905 CCGGAGACGCTCACAGCTTCTGTAGCGGATGCCGGGAGCAGAAAGCCCTCAGGG 3964  
QY 4738 CTATTTTGTCTGCTCTTCTCCGTTATTTAAATTTCTAGTTTGCACCTCCCTTCTGAGAGCAC 4797  
DB |||||  
3965 CGCGTCAAGCGGTTGTGGC----- 3984  
QY 4798 GGCAGTGCAGAGTAGTTAATACTCTGAGGCGAGGCTCTGTGAAGAGTTGCTCTGGGT 4857  
DB |||||  
3985 ----- 3984  
QY 4858 CAGTGTGAGATTTTGGCCATAAAAGGGTCTGCGCCCTGTGTGACAGACAGATCGGAACT 4917  
DB |||||  
3985 ----- 3984  
QY 4918 AGAGTGCATACTCAGAGTCCCGGCTTCCGGGCTCTGATCTCAGGCGATCTTTGCCTA 4977  
DB |||||  
3985 -----GGTGTCCGGGCGCAGCCATGACCCAGTC----- 4012  
QY 4978 GAGATCTCTACGCGGACGATCGTGGCGGGTACCGAGCTCGAATTCGTAATCATGTT 5037  
DB |||||  
4013 -----ACGTAGCGATAGCGGAGTGTATCTGCGTTAACTATGCGGAT 4055  
QY 5038 CATAGCTGTTTCTCTGTGTGAATTTGTTATCGCTCACAATTTCCACACAACATACGAGCG 5097  
DB |||||  
4056 CAGAGCAGATTTGACTGAGAG----- 4076  
QY 5098 GAAGCATAAAGTGTAAAGCCTTGGGGTGCCTAATAGTGAGCTAACTCATAATTTGGT 5157  
DB |||||  
4077 ----- 4076  
QY 5158 TGGCTCACTGCCCGCTTTTCCAGTCGGGAACCTGTGCTGCGCAGCTGCAATTAATGAATCG 5217  
DB |||||  
4077 -----TGACCATAATGCGGTGTAATAATAC 4100  
QY 5218 GCCAACGCGGGGAGAGGGGTTTGGTATTTGGCGCTCTTCCGCTTCTCCGCTCACTG 5277  
DB |||||  
4101 CGCAGATGCTGAAGAGAAAATACCGCATCAGCGCTCTTCCGCTTCTCCGCTCACTG 4160  
QY 5278 ACTCGCTCGCTCGTCTGCTGCTGCGCGGAGCGGTATCAGCTCACTCAAGGCGGTAA 5337









3653	DB	GATTGACGTAGTCGCCCGGGTACCCGGTGATTC	CAATAAACCTCTTGACGTGCAATCCGA	3717
4439	QY	CTCGTGGTCTCGCTGTTCTCTTGGGAGGGTCT	-CTCTGAGTGATGACTACCCCGTCACGG	4497
3713	DB	CTTGTGGTCTCGCTGTTCTCTTGGGAGGGTCT	CTCTGAGTGATGACTACCCGTCACGG	3772
4498	QY	GGTCTTTTCAGTTTCTCCCACTCACAGGTCCT	CACTAACATTCCTGATGTCGCGCAGGA	4557
3773	DB	GGGTCTTTTCATTTGGGGGCTCGTCCGGGAT	CGGAGACCCCTGCCCAGGACCAACCGACC	3832
4558	QY	CTCCGTGAGCCCGGTTTTTGTGTTTATATA	ATAAATCAAGACAGTGTTCCCTTCAAGCCAG	4617
3833	DB	CACCACCGGAGGTAAAGCTG		3852
4618	QY	ACTACATCTGACTCTCGGCTTTATAAAGAA	TGTGAAGGGCTCTGTGACTATCTGCC	4677
3853	DB	-----GCTGCCTCGCGGTTTCGGTGAT	CACGGTGAAACCTCTGAC-ACATGCA	3904
4678	QY	ACAGACTTTTAAAGATTTTATGCCCTCTG	GATGAGGAAATTTAGTCAATCTATCCTCGT	4737
3905	DB	CCCGAGACGGTTCACAGCTTCTCTGTAAG	CGGATCGCGGAGCAGACAAGCCGCTCAGGG	3964
4738	QY	CTATTTTGTGGCTTCTCCGTTATTTAAAT	TTCTAGTTTGCACTCCCTTCTCTGAGGACAC	4797
3965	DB	CGCTCAGCGGTTGTGGCG		3984
4798	QY	GGCGAATTCACAGTAGTTAATACTCTCAG	GGCAGGCTTCTGTGAAGGTTTGCTGGCT	4857
3985	DB	-----	-----	3984
4858	QY	CAGTGTGAGATTTTGCCATAAAGGGGTCT	TGCCCTGTGTACAGACAGATCGGAATCT	4917
3985	DB	-----	-----	3984
4918	QY	AGAGTCATACTCAGAGTCCCGCGGTTCCG	GGGCTCTGATCTCAGGGCATCTTTCGCTA	4977
3985	DB	-----GGTGT	CGGGCGCAGCCATGACCCAGTC	4012
4978	QY	GAGATCCTTACGCCGACGCATCGTGGCG	GGGTACCGAGTTCGTAATCATGTT	5037
4013	DB	-----ACGTAGCGATAGCGGAGTGTA	CTACTGCTTAACTATGCGGCAT	4055
5038	QY	CATAGCTGTTTCTGTGTGAATTTGTTAT	CCGCTCACAAATTCACACAACATACGAGCCG	5097
4056	DB	CAGAGCAGATTGTA	CTGAGAG	4076
5098	QY	GAAGCATAAAGTGTAAAGCCTGGGGT	CCCTAATGATGAGACTAACTCACATTAAT	5157
4077	DB	-----	-----	4076
5158	QY	TGCGCTCACTGCGCGTTTCCAGTTCGG	GAACCTGTCTGTCAGCTGCATTAATGAATCG	5217
4077	DB	-----TGCA	CCATATGCGGTGTGAATATAC	4100
5218	QY	GCCAAACGCGGGGAGAGCGGTTTTCG	TTATGGGCGCTCTTCGCGTTCCTCGCTCACTG	5277
4101	DB	CGCACAGATGCGTAAGGAGAAAAATAC	CGCATCAGGGGCTCTTTCGCTTCCTCGCTCACTG	4160
5278	QY	ACTCGCTGCGCTCGGTGTTTCGCTCG	CGGCGAGCGGTATCAGCTCACTCAAGAGCGCGTAA	5337
4161	DB	ACTCGCTGCGCTCGGTGTTTCGCTCG	CGGCGAGCGGTATCAGCTCACTCAAGAGCGCGTAA	4220
5338	QY	TACGGTTATCCACAGAATCAGGGGATA	AAACGAGAAACATGTGAGCAAAAAGGCCAGC	5397
4221	DB	TACGGTTATCCACAGAATCAGGGGATA	AAACGAGAAACATGTGAGCAAAAAGGCCAGC	4280
5398	QY	AAAAAGGCCAGGAACCGTAAAAAGCC	CGGCTTGCTTCCATGAGCTCCGCGCCCC	5457
4281	DB	AAAAAGGCCAGGAACCGTAAAAAGCC	CGGCTTGCTTCCATGAGCTCCGCGCCCC	4340
5458	QY	CTGACGAGCATCAAAAAATTCGACGCT	CAAGTCAAGGTCGGGAAACCCGACAGGACTAT	5517
4341	DB	CTGACGAGCATCAAAAAATTCGACGCT	CAAGTCAAGGTCGGGAAACCCGACAGGACTAT	4400

Qy	5518	AAAGATACCAAGGGGTTTCCCCTCGGAAAGCTCCCTCGTGCAGCTCTCTCTGTTCCGACCCCTGC	5577
Db	4401	AAAGATACCAAGGGGTTTCCCCTCGGAAAGCTCCCTCGTGCAGCTCTCTCTGTTCCGACCCCTGC	4460
Qy	5578	CGCTTACCGGATACCTGTGCGCGCTTCTCCCTTCGGGAGCGTGGCGCTTCTCATAGCT	5637
Db	4461	CGCTTACCGGATACCTGTGCGCGCTTCTCCCTTCGGGAGCGTGGCGCTTCTCATAGCT	4520
Qy	5638	CACGCTGATGGTATCTCAGTTTCGGTGTAGGTTCGCTTTCGGTAACTATTCGTTTGAAGTCAACAC	5697
Db	4521	CACGCTGATGGTATCTCAGTTTCGGTGTAGGTTCGCTTTCGGTAACTATTCGTTTGAAGTCAACAC	4580
Qy	5698	AAACCCCGCTTACGCCGACCGCTGCGCTTATTCGGTAACTATTCGTTTGAAGTCAACAC	5757
Db	4581	AAACCCCGCTTACGCCGACCGCTGCGCTTATTCGGTAACTATTCGTTTGAAGTCAACAC	4640
Qy	5758	CGGTAAAGACAGACTTATTCGCCACTGGCAGCAGCCACTGGTAAACAGGATTTAGCAGAGCGA	5817
Db	4641	CGGTAAAGACAGACTTATTCGCCACTGGCAGCAGCCACTGGTAAACAGGATTTAGCAGAGCGA	4700
Qy	5818	GGTATGTAGCGGCTGCTACAGAGTTCTTGAAGTGGTGGCTTAACCTACCGCTACACTAGAA	5877
Db	4701	GGTATGTAGCGGCTGCTACAGAGTTCTTGAAGTGGTGGCTTAACCTACCGCTACACTAGAA	4760
Qy	5878	GGACAGTATTTGGTATCTGCGTCTCTGCTGAAGCCAGTTACCTTCGGAAGAGTTGGTA	5937
Db	4761	GGACAGTATTTGGTATCTGCGTCTCTGCTGAAGCCAGTTACCTTCGGAAGAGTTGGTA	4820
Qy	5938	GCTCTTGATCCGGCAACAAACACCGCTGGTGAAGCGGTGGTTTTTTTGGTTTGAAGCAGC	5997
Db	4821	GCTCTTGATCCGGCAACAAACACCGCTGGTGAAGCGGTGGTTTTTTTGGTTTGAAGCAGC	4880
Qy	5998	AGATTACGCGCAGAAAAAGAGTCTCAAGAACATCCTTTGATCTTTCTTACGGGGTCTG	6057
Db	4881	AGATTACGCGCAGAAAAAGAGTCTCAAGAACATCCTTTGATCTTTCTTACGGGGTCTG	4940
Qy	6058	ACGCTCAGTGGAAACGAAACTCACTGTTAAGGATTTTGGTCAATGAGATTATCAAAAGGA	6117
Db	4941	ACGCTCAGTGGAAACGAAACTCACTGTTAAGGATTTTGGTCAATGAGATTATCAAAAGGA	5000
Qy	6118	TCTTCACCTAGATCCCTTTTAAATTAATAATGAAGTTTAAAATCAATCTTAAAGTATATATG	6177
Db	5001	TCTTCACCTAGATCCCTTTTAAATTAATAATGAAGTTTAAAATCAATCTTAAAGTATATATG	5060
Qy	6178	AGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTCAGGAGCCACTTCTCAGCGCATCT	6237
Db	5061	AGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTCAGGAGCCACTTCTCAGCGCATCT	5120
Qy	6238	GTCTATTTGGTTCATCCATAGTTGCCCTGACTCCCGCTCGGTAGATAAACTACGATACGGG	6297
Db	5121	GTCTATTTGGTTCATCCATAGTTGCCCTGACTCCCGCTCGGTAGATAAACTACGATACGGG	5180
Qy	6298	AGGGCTTACCATCTGGCGCCAGTGCTGCAATGATACCGGAGACCCACCGCTCACCGGCTC	6357
Db	5181	AGGGCTTACCATCTGGCGCCAGTGCTGCAATGATACCGGAGACCCACCGCTCACCGGCTC	5240
Qy	6358	CAGATTATCAGCAATAAACCCAGCAGCGGAGGCGCGAGCGCAGAGCTGCTCTGCAA	6417
Db	5241	CAGATTATCAGCAATAAACCCAGCAGCGGAGGCGCGAGCGCAGAGCTGCTCTGCAA	5300
Qy	6418	CTTTATCCGCTCCATCCAGTCTATTAAATTTGTCGGGAAAGCTAGAGTAAGTAGTTCCG	6477
Db	5301	CTTTATCCGCTCCATCCAGTCTATTAAATTTGTCGGGAAAGCTAGAGTAAGTAGTTCCG	5360
Qy	6478	CAGTTAATAGTTTGGCAACGTTTGGTTCGCTTACAGGC-TCGTGGTGTCAAGCTCGT	6536
Db	5361	CAGTTAATAGTTTGGCAACGTTTGGTTCGCTTACAGGC-TCGTGGTGTCAAGCTCGT	5420
Qy	6537	CGTTTGGTATGGCTTCATTACAGTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCC	6596
Db	5421	CGTTTGGTATGGCTTCATTACAGTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCC	5480

6597 CCATGTTGTGCAAAAAGCGGTAGCTCCTTGGTCTCCGATCGTGTGCAAGTAAGT 6656  
 5481 CCATGTTGTGCAAAAAGCGGTAGCTCCTTGGTCTCCGATCGTGTGCAAGTAAGT 5540  
 6657 TGGCCGCACTGTTACACTCATGTTATGGCAGCACTGATTAATCTCTTACTGTCATGC 6716  
 5541 TGGCCGCACTGTTATCCTCATGTTATGGCAGCACTGATTAATCTCTTACTGTCATGC 5600  
 6717 CATCCGTAAGATGCTTTTCTGTGACTGGTGTGAGTACTCAACCAAGTCAATCTCAGATAAGT 6776  
 5601 CATCCGTAAGATGCTTTTCTGTGACTGGTGTGAGTACTCAACCAAGTCAATCTCAGATAAGT 5660  
 6777 GTATGGCGACCGAGTGTCTCTTGGCCGGCGCTCAATACGGGATAATACCGGCCACATA 6836  
 5661 GTATGGCGACCGAGTGTCTCTTGGCCGGCGCTCAACCGGATAATACCGGCCACATA 5720  
 6837 GCAGAACTTTAAAGTGCTCATCTTGGAAAGGTTCTTGGGGGCGGAAACTCTCAGGA 6896  
 5721 GCAGAACTTTAAAGTGCTCATCTTGGAAAGGTTCTTGGGGGCGGAAACTCTCAGGA 5780  
 6897 TCTTACCGCTGTGAGATCCAGTTCGATGTAAACCACTCGTGCACCCCACTGATCTTAC 6956  
 5781 TCTTACCGCTGTGAGATCCAGTTCGATGTAAACCACTCGTGCACCCCACTGATCTTAC 5840  
 6957 CATCTTTTACTTTTACACGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCGCAA 7016  
 5841 CATCTTTTACTTTTACACGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCGCAA 5900  
 7017 AAAAGGAATAAGGGCGACACGGAATGTTGAATCTCATCTCTTCTTTTCAATATT 7076  
 5901 AAAAGGAATAAGGGCGACACGGAATGTTGAATCTCATCTCTTCTTTTCAATATT 5960  
 7077 ATTGAAGCATTTATCAGGGTTATTGCTCATGCGGATACATATTTGAATCTTTTGA 7136  
 5961 ATTGAAGCATTTATCAGGGTTATTGCTCATGCGGATACATATTTGAATCTTTTGA 6020  
 7137 AAAATAAAACAATAGGGTTCCGCGCACATTTCCCGGAAAAGTGCCACCTGACGCTTAAG 7196  
 6021 AAAATAAAACAATAGGGTTCCGCGCACATTTCCCGGAAAAGTGCCACCTGACGCTTAAG 6080  
 7197 AAACCATTTATATGACATTAACCTATAAAAATAGGCGTATACGAGGCCCTTTGTC 7256  
 6081 AAACCATTTATATGACATTAACCTATAAAAATAGGCGTATACGAGGCCCTTTGTC 6140  
 7257 T 7257  
 6141 T 6141

RESULT 8  
 US-08-848-760B-1  
 ; Sequence 1, Application US/08848760B  
 ; Patent No. 6248721  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chang, Lung-Ji  
 ; TITLE OF INVENTION: Animal Model For Evaluation Of Vaccines  
 ; NUMBER OF SEQUENCES: 33  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESS: Saliwanchik, Lloyd & Saliwanchik  
 ; STREET: 2421 N.W. 41st Street, Suite A-1  
 ; CITY: Gainesville  
 ; STATE: Florida  
 ; COUNTRY: United States of America  
 ; ZIP: 32606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patencin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/848,760B  
 ; FILING DATE: 25-Jan-2001  
 ; CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/838,702  
 FILING DATE: 09-APR-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: PACE, DORAN R.  
 REGISTRATION NUMBER: 38,261  
 REFERENCE/DOCKET NUMBER: CNG-100C1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (352) 375-8100  
 TELEFAX: (352) 372-5800  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 6145 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-08-848-760B-1  
 Query Match 40.0%; Score 3062.2; DB 3; Length 6145;  
 Best Local Similarity 71.7%; Pred. No. 0;  
 Matches 4818; Conservative 0; Mismatches 1168; Indels 735; Gaps 32;  
 QY 549 TTTTGAAGAGCCCACTGTAGTTTGGCAAGCTAGCTTAAAGTAACGCCATTTTCAAGG 608  
 DB 144 TTTTGAAGAGCCCACTGTAGTTTGGCAAGCTAGCTTAAAGTAACGCCATTTTCAAGG 201  
 QY 609 CATGG-AAAATACATACTGAGATAGAGAGTTTCAGATCAAGTTTAGGAAACAGAGAC 667  
 DB 202 CATGAAAAATACATACTGAGATAGAAAAAGTTTCAGATCAAGTTTAGGAAACAGAGAC 261  
 QY 668 AGCAATATGGGCCCAACAGAGATATCTGTGTAGAGCTTCTGCCCC-GCTCAGGGCC 726  
 DB 262 AGCTGAATA-CCAAACAGAGATATCTGTGTAGAGCTTCTGCCCCGGCTCAGGGCC 317  
 QY 727 AAGAACAGTTGGAAACAGAGAGAT-ATGGGCCAAACAGGATATCTGTGTAGAGCTTCT 785  
 DB 318 AAGAACAGATGAGACAGCTGAGTGATGGCCCAACAGGATATCTGTGTAGAGCTTCT 377  
 QY 786 GCCCGGCTCAGGGCCAAAGAACAGATGTCCTCAGATGGGTCCCGCTCAGAGATTTTC 845  
 DB 378 GCCCGGCTCAGGGCCAAAGAACAGATGTCCTCAGATGGGTCCCGCTCAGAGATTTTC 437  
 QY 846 TAGAACAATCAGATGTTTCCAGGTGCCCCAGGACCTG-AAATGACCTGTGCTTA 904  
 DB 438 TAGTAATCAGATGTTTCCAGGTGCCCCAGGACCTGAAATGACCTGTGCTTA 497  
 QY 905 TTTGAACTAAACCAATCAGTTGCTTCTGCTTCTGCTGCGCGCTTCTGCTCCCGAGCT 964  
 DB 498 TTTGAACTAAACCAATCAGTTGCTTCTGCTTCTGCTTCTGCTGCGCGCTTCTGCTCCGAGCT 557  
 QY 965 CAATAAAGAGCCCAACCCCTCACTCGGCGCGCAGTCTCTCCGATAGACTGCGTCCGC 1024  
 DB 558 CAATAAAGAGCCCAACCCCTCACTCGGCGCGCAGTCTCTCCGATAGACTGCGTCCGC 617  
 QY 1025 CGGTACCCGTTATCCCAATAAAGCTCTTGTGTTTGCATCCGATCCGATCGTGCATCGCTG 1084  
 DB 618 CGGTACCCGTTATCCCAATAAAGCTCTTGTGTTTGCATCCGATCCGATCGTGCATCGCTG 677  
 QY 1085 ATCTTTGGAGGGTCTCTCTCAGATTGATTGACTGCGCCACCTCGGGGGTCTTTCAATTGA 1144  
 DB 678 TTCTTTGGAGGGTCTCTCTCAGATTGATTGACTGCGCCAGGACCTGAAATGACCTTTCAATTGGG 737  
 QY 1145 GGTTCACCGAGATTGGAGACCCCTGCCAGGACCAACCGACCCCGCGCGGAGTA 1204  
 DB 738 GGTTCGTCCGGATTGGAGACCCCTGCCAGGACCAACCGA-CCCAACCGCGGAGTA 796  
 QY 1205 AGTGGCAGCACTTATCTGTCTGTCCGATGTTGTCTAGTGTCTATGATGATTTATG 1264  
 DB 797 AGTGGCAGCACTTATCTGTCTGTCCGATGTTGTCTAGTGTCTATGTTGATGTTATG 856  
 QY 1265 GCCTGTGCGTACTAGTACTAGTAACTAGCTCTGTATCTGCGGACCCGCTGGTGAAC 1324







QY	3482	ATAACGAGCTCAATCTAGACGAAAGAGAGAGTACGATGTTTGGACAAAGACGCTGGCC	3541	QY	4558	CTCCGTCAGCCCGGTTTTTGTATTATAATAAATGAAGAACAGTGTTCCTTTCAAGCCAG	4617
Db	2823	AGAACGAATGCTTGCAGGCATCTCATGATGTGTCTT-----CCCGTTTC	2869	Db	3833	CACCAACCCGGAGGTAAGCTG-----	3852
QY	3542	GGGACCTCAGATGGGGGAAAGCCGAGAGAAACCCCTCAGGAAGCCCTGTACAATG	3601	QY	4618	ACTACATCTGACTCTCGGCTTTATAAAGAAATGTTGAAGGCTCTGTGGACTATCTGCC	4677
Db	2870	CGCCTGAGTCACTGCGTGGATGAGCGCTGCGCCTGCTGCGCGAGCGGAG---CTG	2925	Db	3853	-----GCTGCTCGCGGTTTCGGTGATGACGGTGAAACCTCTGAC-ACATCAGCT	3904
QY	3602	AACTGCAGAAAGTAAGATGGCGGAGGCTTACAGTGAGATTGGGATGAAGCGAGCGCC	3661	QY	4678	ACACGACTTTTTAAGATTTTTATGCTCTGATGAGGATTTTGTCAATCTATCTCTGT	4737
Db	2926	CTCACACCCACTCGAGGCGCTGACGCTGCAGAGCGCGAGTCAGAACTGC-----	2978	Db	3905	CCCGAGAGCGGTACAGCTTGTCTGAAGCGATCCCGGAGCAGACAGCCCGTCAGGG	3964
QY	3662	GGAGGGCAAGGGGACGATGCTTTTACAGGGTCTCAGTACAGCCACCAAGGACACCT	3721	QY	4738	CTATTTGCTGCTCTTCTCCGTATTTTAAATTTCTAGTTTGCACCTCCCTTCTCTGAGAGCAC	4797
Db	2979	-----TCCAAAGGACCTCAAGGCTTTCGAGGGACACTAGGCTGACTCATTCGAGCCAG	3033	Db	3965	CGCGTCAGCGGCTGTGGCG-----	3984
QY	3722	ACGAGCCCTTACATGACAGGCCCTGCCCTCGCTAACTGACGCGGCGCGGATCCGG	3781	QY	4798	GGCGATTGCAGAGTAGTTAACTCTGAGGGCAGGCTTCTGTGAAGAGGTTCCTTGGGCT	4857
Db	3034	TGTA-----GAGATAAGCTTATCG	3052	Db	3985	-----	3984
QY	3782	ATTAGTCCAAATTGTTAAGACAGGATATCAGTGTTCAGGGCTCTAGTTTTCAGTCAACA	3841	QY	4858	CAGTGTGAGATTTTCCCATATAAAGGGGTCTTGCCCTGTGTGACAGACAGATCGGAATCT	4917
Db	3053	ATTAGTCCAAATTGTTAAGACAGGATATCAGTGTTCAGGGCTCTAGTTTTCAGTCAACA	3112	Db	3985	-----	3984
QY	3842	ATATCACAGCTGAAGCCTATAGATACGAGCCATAGATAAAATAAAGATTTTATTAG	3901	QY	4918	AGAGTGCATACTCAGAGTCCCGCGGTTCCCGGGCTCTGATCTCAGGGCATCTTTGCCCTA	4977
Db	3113	ATATCACAGCTGAAGCCTATAGATACGAGCCATAGATAAATAAAGATTTTATTAG	3172	Db	3985	-----GGTCTCGGGCGCAGCCATGACCCAGTC-----	4012
QY	3902	TCTCCAGAAAAAGGGGGAAAGAACCCACCTGTAGTTTGGCAAGCTAGCTTAAAT	3961	QY	4978	GAGATCTCTACGCCGAGACGATCGTGGCGGGTACCGAGCTCGAATTCGTAATCATGCT	5037
Db	3173	TCTCCAGAAAAAGGGGGAAAGAACCCACCTGTAGTTTGGCAAGCTAGCTTAAAT	3232	Db	4013	-----ACGTAGCCATAGCGAGTGTATCTGCTTAACTATGCGCAT	4055
QY	3962	AAAGCCATTTGCAAGGCTAGG-AAAAATACATACTGAGATAGAGAAGTTCCAGATCAAG	4020	QY	5038	CATAGCTGTTTCTGTGTGAATTTGTATCCGCTCACAAATTCACACACACATACGAGCG	5097
Db	3233	AAAGCCATTTGCAAGGCTAGGAAAAATACATACTGAGATAGAGAAGTTCCAGATCAAG	3292	Db	4056	CAGACGATTTGTACTGAGAG-----	4076
QY	4021	GTTAGGAACAGA-GAGACAGCAGATATGCGGCCAACAGGATATCTGTGTGAAGCATTC	4079	QY	5098	GAAACATAAAGTGTAAAGCTGGGGTGCTTAATGAGTGAGCTAACTCACATTTAATTTGCT	5157
Db	3293	GTCAGGAACAGATGMAACAGCTGAATATGCGGCCAACAGGATATCTGTGTGAAGCATTC	3352	Db	4077	-----	4076
QY	4080	CTGCCCC-GCTCAGGCCCAAGAACAGTTGGAAACAGAGATATGCGGCCAACAGGATATC	4138	QY	5158	TGCGCTACTGCCCGCTTTCCAGTCGGGAAACCTGTCTGTGCGAGCTGCATTAATGAATCG	5217
Db	3353	CTGCCCCGCTCAGGCCCAAGAACAGTGAACAGCTGAATATGGGCCAACAGGATATC	3412	Db	4077	-----TGACCATATGCGGTGTGAATAC	4100
QY	4139	TGTGTAAGCAGTTCTGCCCCGGCTCAGGGCCAAAGAACAGATGTTCCAGATGGGTC	4198	QY	5218	GCCAAACCGCGGGAGAGCGGTTTGCATATTGGCGCTCTTCCGCTTCTCGCTCAGCTG	5277
Db	3413	TGTGTAAGCAGTTCTGCCCCGGCTCAGGGCCAAAGAACAGATGTTCCAGATGGGTC	3472	Db	4101	CGCAGATGCTGAAGAGAAATAACCGCATACGCGCTCTTCCGCTTCTCGCTCAGCTG	4160
QY	4199	CGCCCTCAGCAGTTCTTAGAACCATCAGATGTTTCCAGGGTGCCCCCAAGGACTGAA	4258	QY	5278	ACTCGCTCGCTCGGTCGTTCCGCTGCGGCGAGCGGTATCAGCTCACTCAAGAGCGGTAA	5337
Db	3473	CAGCCCTCAGCAGTTCTTAGAACCATCAGATGTTTCCAGGGTGCCCCCAAGGACTGAA	3532	Db	4161	ACTCGCTCGCTCGGTCGTTCCGCTGCGGCGAGCGGTATCAGCTCACTCAAGAGCGGTAA	4220
QY	4259	ATGACCTGTGCTTATTGAACTAACCAATCAGTTCTGCTCTCGCTTCTGTTCCGCGC	4318	QY	5338	TACGTTTATCCAGAAATCAGGGGATAACGCAAGAAAGAACATGTGAGCAAAAGGCCAGC	5397
Db	3533	ATGACCTGTGCTTATTGAACTAACCAATCAGTTCTGCTCTCGCTTCTGTTCCGCGC	3592	Db	4221	TAGGTTTATCCAGAAATCAGGGGATAACGCAAGAAAGAACATGTGAGCAAAAGGCCAGC	4280
QY	4319	TTCGTCTCCCGAGCTCAATAAAGAGCCCAAAACCCCTCACTCGGCGGCCAGTCTCC	4378	QY	5398	AAAAGCCAGGAACCGTAAAAAGCGGTTGCTGCGGCTTTTCCATAGGCTCCGCCCC	5457
Db	3593	TTCGTCTCCCGAGCTCAATAAAGAGCCCAAAACCCCTCACTCGGCGGCCAGTCTCC	3652	Db	4281	AAAAGCCAGGAACCGTAAAAAGCGGTTGCTGCGGCTTTTCCATAGGCTCCGCCCC	4340
QY	4379	GATAGCTCGTTCGCGCGGTACCGGTGTTCTCAATTAACCTCTTGCATTCGATCCGA	4438	QY	5458	CTGACGACATCACAAAAATCAGCTCAAGTTCAGAGTGGGGAACCCGACAGGACTAT	5517
Db	3653	GATGACTGAGTCGCGCGGTACCGGTGTTCTCAATTAACCTCTTGCATTCGATCCGA	3712	Db	4341	CTGACGACATCACAAAAATCAGCTCAAGTTCAGAGTGGGGAACCCGACAGGACTAT	4400
QY	4439	CTCGTGTCTGCTGTTCTTGGAGGGTCT-CTCTGAGTGAATGACTACCCGTCAGCGG	4497	QY	5518	AAAGTACACAGCGTTTCCCTCGTGAAGCTCCCTCGTGGCTCTCTCTGTTCGACCCCTGC	5577
Db	3713	CTTGTGTCTGCTGTTCTTGGAGGGTCTCTCTGAGTGAATGACTACCCGTCAGCGG	3772	Db	4401	AAAGTACACAGCGTTTCCCTCGTGAAGCTCCCTCGTGGCTCTCTCTGTTCGACCCCTGC	4460
QY	4498	GCTCTTTCAGTTTCTCCACCTACACAGTCTCACTAAATTTCTGATGTGCGCAGGGA	4557	QY	5578	CGCTTACCGGATACTGTCTCGGCTTTTCCCTTCCGGGAAGCGTGGCGCTTTCTCATAGCT	5637
Db	3773	GGTCTCTTCAATTTGGGGGCTCGTCCGGGATCGGAGACCCCTCCCGCAGGACCAACGACC	3832	Db	4461	CGCTTACCGGATACTGTCTCGGCTTTTCCCTTCCGGGAAGCGTGGCGCTTTCTCATAGCT	4520
				QY	5638	CACGCTGTAGGTATCTCAGTTTCGGTGTAGTTCGTTTCGCTCCAAAGCTGGGCTGTGTGCAG	5697



;	INFORMATION FOR SEQ ID NO: 1:	
;	SEQUENCE CHARACTERISTICS:	
;	LENGTH: 6145 base pairs	
;	TYPE: nucleic acid	
;	STRANDEDNESS: double	
;	TOPOLOGY: linear	
;	MOLECULE TYPE: DNA (genomic)	
;	SEQUENCE DESCRIPTION: SEQ ID NO: 1:	
;	US-09-826-025-1	
	Query Match	40.0%; Score 3062.2; DB 4; Length 6145;
	Best Local Similarity	71.7%; Pred. No. 0;
	Matches 4818; Conservative	0; Mismatches 1168; Indels 735; Gaps 32;
Qy	549	TTTTGAAAGACCCACCTGTAGGTTTCGGCAAGCTAGCTTAAGTAACGCCATTTTGCAAGG 608
Db	144	TTTTGAAAGACCCACCCCGTAGG--TGGCAAGCTAGCTTAAGTAACGCCACTTTCGAGG 201
Qy	609	CATGG-AAAATACATAA CTGAGATAGAGAAGTTTCAGATCAAGTTTAGGAACAGAGAC 667
Db	202	CATGGAATAATACATAA CTGAGAATACAAAGTTTCAGATCAAGTCAAGAACAAAGAAAC 261
Qy	668	AGCAGATATAGGGCAACAGGATATCTGTGTTAAGCAGTTCTCTGCCCC--GCTCAGGGCC 726
Db	262	AGCTGAATA- ---CCAAACAGGATATCTGTGTTAAGCGGTTCTTGCCCCCGGCTCAGGGCC 317
Qy	727	AAGAACAGTTTGGAAACAGAGAAAT-ATGGGCAAAACAGGATATCTGTGTTAAGCAGTTCC 785
Db	318	AAGAACAGATGAGACACTGAGTGATGGGCCAAACAGGATATCTGTGTTAAGCAGTTCC 377
Qy	786	GCCCCGGCTCAGGSCCAAGAACAGATGGTCCCCAGATGCGGTCCCGCCCTCAGCAGTTTC 845
Db	378	GCCCCGGCTCGGGSCCAAGAACAGATGGTCCCCAGATGCGGTCCAGCCCTCAGCAGTTTC 437
Qy	846	TAGAAACCATCAGATGTTTCAGGGGTGCCCAAGGACCTG-AAATGACCTGTGCCTTA 904
Db	438	TAGTGAATCATCAGATGTTTCCAGGGTGCCCCAAGGACCTGAAATGACCTGTGACCTTA 497
Qy	905	TTTGAACCTAACCAATCAGTTTCGCTTTCGCTTTCGTTTCGTCGCGCTCTGTGTCGCCAGCT 964
Db	498	TTTGAACCTAACCAATCAGTTTCGCTTTCGCTTTCGTTTCGTCGCGCTTCCGCTCTCCGAGCT 557
Qy	965	CAATAAAGAGCCCAACAACCCCTCACCTCGGCGGCCAGTCTCTCGATAGACTGCGTCGCC 1024
Db	558	CAATAAAGAGCCCAACAACCCCTCACCTCGGCGGCCAGTCTCTCGATAGACTGCGTCGCC 617
Qy	1025	CGGGTACCCGTAATTCACCAATAAGACCTCTGCTGTTTGGCATCCGAATCGGTGAGACTGCGTG 1084
Db	618	CGGGTACCCGTAATTCACCAATAAGACCTCTGCTGTTTGGCATCCGAATCGGTGCTCGCTG 677
Qy	1085	ATCCTTGGGAGGGTCTCCTCAGATTGATGCACTGCCACCTCGGGGTCTTTCATTTGGA 1144
Db	678	TTCTCTGGGAGGGTCTCCTCTGAGTGATGACTACCCACGACGGGGGTCTTTCATTTGGG 737
Qy	1145	GGTTTCCACCGAGATTTGGAGACCCCTGCCAGGGACCAACCGACCCCCCGCCGGAGGTA 1204
Db	738	GGCTCGTCCGGATTTGGAGACCCCTGCCACGAGACCAACGA-CCACACACCCGGAGGTA 796
Qy	1205	AGCTGGCCAGCAACTTATCTGTGTCGTGTCGATTTGCTAGTGTCTATGACTGATTTTATG 1264
Db	797	AGCTGGCCAGCAACTTATCTGTGTCGTGTCGATTTGCTAGTGTCTATGTTTGTATGTTATG 856
Qy	1265	CGCCTGCGTCCGTACTAGTAACTAGCTAACTAGCTCTGTATCTGGCGGACCGGTGGTGA 1324
Db	857	CGCCTGCGTCTGTACTAGTTAGCTTAACTAGCTCTGTATCTGGCGGACCCGTGGTGGAACT 916
Qy	1325	GACGAGTTTCGGAACAACCCCGGCCGCAACCTCTGGGAGAGCTCCAGGAGACTTCGGGGGCGGT 1384
Db	917	GACGAGTTCTGAACACACCCGGCCGCAACCTCTGGGAGAGCTCCAGGAGACTTTGGGGGCGGT 976
Qy	1385	TTTTGTGGCCGCACTGAGTCTTAAATCCGATTCGTTTATGAGACTCTTTGTGTGCAACCCCC 1444
Db	977	TTTTGTGGCCGCACTGAGGAAGGGAGTTCGATGTGGAAATCCGACCCC----- 1022

QY 2522 CCAGGTGGGTGACAGAGTGACCATCACCTGTAGGCCAGTTCAGATGTGGTACTTCTG 2581  
Db 2012 AATCGCGGGTGCATACGCTGTATCCGGCTACCTGCCCATTCGACCAACCAAGCGAAACA 2071  
QY 2582 TAGCTTGTGACAGCAGAGGACAGGAGTGAAGGCTCCAAAGCTGCTGATCTACTCTGAGACATCCA 2641  
Db 2072 TCGCATCGAGCGAGCACGCTACTCGGATGAGAGCGGCTCTGTCTGATCAGGATGATCTGGA 2131  
QY 2642 CCGGCGACACTGTGTGTCAGAGCAGATTCAAGCGGTAGCGGTAGCGGTACCGACTTCACCT 2701  
Db 2132 CGAAGAGCATCAGGGGCTCGCGCCAGCGCAACTGT-----TCGCGCAGGC 2175  
QY 2702 TCACCATCAGCAGCTCCAGCCAGGACATCGCCACCTACTACTGCGCAGCAATATAGCC 2761  
Db 2176 TCAAGGCGCATGCGCCGACGCGGAGGATCTCGTGTGACCCATGGCGATGCCCTGCTGCG 2235  
QY 2762 TCTATCGGCTCGTTCGGCCAGGAGCAAGGTGGAATCAAAAGAGGTGGCTCAGGATCGG 2821  
Db 2236 CGAATATCATGTTGGAAATGCGCGCTTTCTGGATTCATCGACTGTGGCCGGCTGGGTG 2295  
QY 2822 GTGATCCGGCTCTGTGGCTCAGGATCGGAGGTCCAACTGGTGGAGAGCGGTGGAGGTG 2881  
Db 2296 TGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCCGTGATATTGCTGAAGAGCTTGGCG 2355  
QY 2882 TTGTGCAACCTGGCGGCTCCCTGCGCCTGTCTGCTCGCATCTGGCTTCGATTCACCA 2941  
Db 2356 CGAATCGGCTGACCGCTTCTCGTGTCTTACGGTATCGCGCTCCC----- 2402  
QY 2942 CATATTGATGATGGGTGAGAGCAGGACCTGGAAAGGTCTTGAGTGGATTGGAGAA 3001  
Db 2403 -----GATTCGAGCGCATCGCCTTCTATCGCCTTCTTGACGAG 2441  
QY 3002 TTCAATCCAGATPAGCAGTACGATTAATATGCGCGTCTCTAAGGATAGATTTACAATAT 3061  
Db 2442 TTCTTCTGAGCGGACTCTGCGGTTCGAAATGACCGACCAAGCGACGCCCAACCTGCCAT 2501  
QY 3062 CGCGAGCAACGCCAAGNACATTTGCTCGCAATGGACAGCCTGAGACCCCGAGACA 3121  
Db 2502 CACGAGATTTTCGATTCACCCCGCCTTCTATGAAAGTTGGGCTTCGGAATCGTTTTCC 2561  
QY 3122 CCGGGTCTATTTTGTGCAAGCCTTTACTTCGGCTTCCCTGTGTTGCTTATTGGGGCC 3181  
Db 2562 GGGAGGCC-----GGTGGATGATCCT 2583  
QY 3182 AAGGAGCCCCGGTCAACCGTCTCCAGTGTAAAGCCCAACGACGCGCAGCGCGGACAC 3241  
Db 2584 CCAGCGGGGATCTCATGCTGGAGTCTTTCGCCACCCCGGGCTCGATCCCTCGCGAG 2643  
QY 3242 CAACACCGGCGCCACCATCGGTGCGAGCCCTGTCTCGGCCGAGGCGGCTCGGC 3301  
Db 2644 TTGGTTCACTGTGCTGCTGAGGCTGGACGACTCG----- 2678  
QY 3302 CAGCGGGGGGGCGCAGTGACACGAGGGGCTGAGCTTCGCTCGCATCCCAACTCT 3361  
Db 2679 -----CGAGTTCTTACCGCAGGTGCAATCCG 2705  
QY 3362 GCTACCTGCTGGATGGAATCTCTTCTATGTGTGTCATTCATCTACCTGCTTGTCTCTGA 3421  
Db 2706 TCGGCATCCAGGAA--AACACGACGCGGCTATCCGCGCATCCATGCCCGCCGAACTGCAGG 2762  
QY 3422 GAGTGAAGTTGACGAGAGCGAGAGCCCCCGGTATCAGAGGGCCAGAACAGCTCT 3481  
Db 2763 AGTGGGAGGACGATGTCGCTTTGTGTCGACCCCGACCGGACGCTCTCGCGCTGTATAC 2822  
QY 3482 ATACGAGCTCAATCTAGACGAGAGAGGAGTACGATGTTTGGACAAGAGAGCTGGCC 3541  
Db 2823 AGAACGAATGCTGACGGCATCTCATGAGTGTGTCTT-----CCCGTTTTTC 2869  
QY 3542 GGGACCTTGAGATGGGGGAAAGCGAGAGGAAGAACCTTCAGGAAGCGCTGTACAAATG 3601  
Db 2870 CGCTGAGTCACTGCGTGGATGAGCGCTGGCGCTCTGCGCGACGCGCAG-----CTG 2925  
QY 3602 AACTGCAGAAAGATAAGATGGCGGAGCGCTTACAGTGAGATTGGGATGAAAGCGCGGCC 3661

Db 2926 CTACACCCACTCGAGGGCGTGCAGGCGTGCAGAGCCGAGTGCAGAACTGC----- 2978  
QY 3662 GGAGGGGCAAGGGGACAGATGGCCCTTTTACCAGGGTCTCAGTACAGCCACCAAGGACACT 3721  
Db 2979 -----TCCAAAGGGAGCTCAAGGCTTTCCGAGGGACACTAGGCTGACTCCATCGAGCCAG 3033  
QY 3722 ACAGAGCCCTTTCATCGAGCCCTGCCCCCTCTCTAACTCGACGCGCGCGGATCCGG 3781  
Db 3034 TGTA-----GAGATAAGCTTATCG 3052  
QY 3782 ATTAGTCCAAATTTCTTAAAGACAGGATATCAGTGTCCAGGCTCTAGTTTTGACTCAACA 3841  
Db 3053 ATTAGTCCAAATTTCTTAAAGACAGGATATCAGTGTCCAGGCTCTAGTTTTGACTCAACA 3112  
QY 3842 ATATCACCAGCTGAAGCTTATAGATGACGACCATAGATAAAATAAAGATTTTATTTAG 3901  
Db 3113 ATATCACCAGCTGAAGCTTATAGATGACGACCATAGATAAAATAAAGATTTTATTTAG 3172  
QY 3902 TCTCCAGAAAAGGGGGGAATGAAAGACCCCACTGTAGGTTTGGCAAGCTAGCTTAAGT 3961  
Db 3173 TCTCCAGAAAAGGGGGGAATGAAAGACCCCACTGTAGGTTTGGCAAGCTAGCTTAAGT 3232  
QY 3962 AACGCCATTTTGCAGGATCG-AAAATACATACACTCAGATAGAGAGTTTCAGATCAAG 4020  
Db 3233 AACGCCATTTTGCAGGATCGAAAAATACATACTGAGATAGAGAGTTTCAGATCAAG 3292  
QY 4021 GTTAGGAACAGA--GAGACAGCAGCAATATGGGCCAAAACAGGATATCTGTGTAAAGCAGTTC 4079  
Db 3293 GTCAGGAACAGATGGAACAGCTGAATATGGGCCAAAACAGGATATCTGTGTAAAGCAGTTC 3352  
QY 4080 CTGCCCC--GCTCAGGGGCAAGAAACAGTTGGAACAGGAGATATATGGGCCAAAACAGGATATC 4138  
Db 3353 CTGCCCCGGCTCAGGGGCAAGAAACAGATGGAACAGCTGAATATATGGGCCAAAACAGGATATC 3412  
QY 4139 TGTGTTAAGCAGTCTTCTGCCCCGGCTCAGGGCCAAAGAACAGATGTTCCCAAGATGCGGTG 4198  
Db 3413 TGTGTTAAGCAGTCTTCTGCCCCGGCTCAGGGCCAAAGAACAGATGTTCCCAAGATGCGGTG 3472  
QY 4199 CCGCCCTCAGCAGTCTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCCCAAAGGACCTGAA 4258  
Db 3473 CAGCCCTCAGCAGTCTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCCCAAAGGACCTGAA 3532  
QY 4259 ATGACCTGTGCTTATTTGAACTAACTAAGTTCAGTTCGCTTCTGCTTCTGTTGCGCGGC 4318  
Db 3533 ATGACCTGTGCTTATTTGAACTAACTAAGTTCAGTTCGCTTCTGCTTCTGTTGCGCGGC 3592  
QY 4319 TTCTGCTCCCGAGCTCAATAAAGAGCGCACACCCCTCACTCGGCGCGCAGTCTCTCC 4378  
Db 3593 TTCTGCTCCCGAGCTCAATAAAGAGCGCACACCCCTCACTCGGCGCGCAGTCTCTCC 3652  
QY 4379 GATAGACTGCTGCTCCCGGTACCCGTGTTCTCAATAAACCCCTCTTGAGTTGCATCCGA 4438  
Db 3653 GATTGACTGAGTCCCGGGTACCCGTGTATCCAAATAAACCCCTCTTGAGTTGCATCCGA 3712  
QY 4439 CTGCTGCTGCTGCTTCTTGGAGGAGTCT-CTCTGAGTGAATGACTACCCGTCAGCGG 4497  
Db 3713 CTGCTGCTGCTGCTTCTTGGAGGAGTCTCTCTGAGTGAATGACTACCCGTCAGCGG 3772  
QY 4498 GGTCTTTTCACTTCTCCCACTACAGGTTCTCAATCACTCTGATGTGCGCGCAGGGA 4557  
Db 3773 GGGTCTTTTCACTTGGGGGCTCGTCCGGATCGGAGACCCCTGCCAGGGACCCAGCC 3832  
QY 4558 CTCCGTGAGCCCGGTTTTTTTATATAATAAATGCAAGAACAGTGTTCCTTCAAGCCAG 4617  
Db 3833 CACCACCGGAGGTAAAGCTG----- 3852  
QY 4618 ACTACATCTGACTCTCGGCTTTTATAAAGAAATGTTGAAGGGCTCTGTGAGTATCTGCC 4677  
Db 3853 -----GCTGCCTCGCGGTTTTCGGTGATGACGGTGAACACCTCTGAC-ACATGCACT 3904  
QY 4678 ACAGACTTTTAAAGATTTTATGCCCTCTCGGTGAGGAGTTTGTCTATCTATCTCTCTGT 4737

Db 3905 CCCGAGACGGTACACAGCTTGTCTGTAAAGCGGATGCCGGAGCAGACAAGCCCGTCAGGG 3964  
Qy 4738 CTATTTGCTGGCTTCTCGTATTTTAAATTTCTAGTTTGCATCCCTTCTGAGAGCAC 4797  
Db 3965 CCGCTCAGCGGGTGTGGCG----- 3984  
Qy 4798 GCGGATGACAGAGTAGTTAATACTCTGAGGCGAGGCTTCTGTGAAAAGTGTGCTGGGCT 4857  
Db 3985 ----- 3984  
Qy 4858 CAGGTGAGATTTTGCCATAAAAAAGGGTCTTGCCCCCTGTGTACAGACAGATCGGAATCT 4917  
Db 3985 ----- 3984  
Qy 4918 AGAGTGCATACTCAGAGTCCCCCGGTTCCGGGGCTCTGATCTCAGGGCATCTTTGCCTA 4977  
Db 3985 -----GGTGTCCGGGGCGCAGCCATGACCCAGTC----- 4012  
Qy 4978 GAGATCCTCTAGCCCGGAGCGCATCTGTGGCCGGGTACCGAGCTCGAATTCGTAATCATGGT 5037  
Db 4013 -----ACGTAGCGATAGCGGAGTGATCTGGCTTAATATGGGCT 4055  
Qy 5038 CATAGCTGTTTCCGTGTGAAATGTGTATCCGCTCACAAATCCACACAATACGAGCGG 5097  
Db 4056 CAGAGCAGATTGTACTGAGAG----- 4076  
Qy 5098 GAAGCATAAAGTGTAAAGCTCGGGTGCCTAATGAGTAGCTAACTCACATTAATTTGGT 5157  
Db 4077 ----- 4076  
Qy 5158 TGGCTCACTGCCCGCTTCCAGTCGGGAAACCTGTGTGCCAGCTGCAATTAATGAATCG 5217  
Db 4077 -----TGACCATATGCGGTGTGAATAC 4100  
Qy 5218 GCCAAGCGCGGGAGAGCGGTTTTCGTAATTTGGGGCTTCTTCGCTTCCGCTCACTG 5277  
Db 4101 CGCACAGATGCGTAAGGAGAAATACCGCATCAGGCGCTTCTTCGCTTCCGCTCACTG 4160  
Qy 5278 ACTCGCTGCGCTCGGTGCTTCCGCTCGCGGAGCGGTATCAGCTCACTCAAAAGGGGTAA 5337  
Db 4161 ACTCGTGGCGTCCGTGCTTCCGCTGCGCGAGCGGTATCAGCTCACTCAAAAGGGGTAA 4220  
Qy 5338 TACGGTTATCCACAGAAATCAGGGGATAACCGAGGAAAGAAATGTGAGCAAAAGGCCAGC 5397  
Db 4221 TACGGTTATCCACAGAAATCAGGGGATAACCGAGGAAAGAAATGTGAGCAAAAGGCCAGC 4280  
Qy 5398 AAAAGGCCAGGAACCGTAAAAAGGCGGCTTGTGCGCTTTTCCATAGGCTCCGCCCC 5457  
Db 4281 AAAAGGCCAGGAACCGTAAAAAGGCGGCTTGTGCGCTTTTCCATAGGCTCCGCCCC 4340  
Qy 5458 CTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTAT 5517  
Db 4341 CTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTAT 4400  
Qy 5518 AAAAGTACCAGCGGTTTCCCTCGAAGCTCCCTCGTGGCTCTCTCTGTTCGACCCCTGC 5577  
Db 4401 AAAAGTACCAGCGGTTTCCCTCGAAGCTCCCTCGTGGCTCTCTCTGTTCGACCCCTGC 4460  
Qy 5578 CCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCTATAGCT 5637  
Db 4461 CCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCTATAGCT 4520  
Qy 5638 CAGCGTGTAGGTATCTCAGTTCCGTTAGGTGTGTTCCGCTCCAGCTGGGCTGTGTGCAGC 5697  
Db 4521 CAGCGTGTAGGTATCTCAGTTCCGTTAGGTGTGTTCCGCTCCAGCTGGGCTGTGTGCAGC 4580  
Qy 5698 AACCCCGCTTACGCGCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAAC 5757  
Db 4581 AACCCCGCTTACGCGCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAAC 4640  
Qy 5758 CCGTAAACACGACTTATCCGCTACGCGGAGCGGCTGCTGTTAAAGGATTTAGCAGGCGA 5817  
Db 4641 CCGTAAACACGACTTATCCGCTACGCGGAGCGGCTGCTGTTAAAGGATTTAGCAGGCGA 4700

Qy 5818 GGTATGTAGGCGGTCTACAGAGTTCTTGAAGTGTGGCTAACTACCGCTACACTAGAA 5877  
Db 4701 GGTATGTAGGCGGTCTACAGAGTTCTTGAAGTGTGGCTAACTACCGCTACACTAGAA 4760  
Qy 5878 GGACAGTATTTGGTATCTGCGCTTGTCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGTA 5937  
Db 4761 GGACAGTATTTGGTATCTGCGCTTGTCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGTA 4820  
Qy 5938 GCTCTTGATCCGGGCAAAACAAACCCAGCTGGTAGCGGTGTTTTTTTGTTCGACGAC 5997  
Db 4821 GCTCTTGATCCGGGCAAAACAAACCCAGCTGGTAGCGGTGTTTTTTTGTTCGACGAC 4880  
Qy 5998 AGATTACCGGCAAAAAAGGATCTCAAGAGATCTCTTTGATCTTTTCTACGGGTCTG 6057  
Db 4881 AGATTACCGGCAAAAAAGGATCTCAAGAGATCTCTTTGATCTTTTCTACGGGTCTG 4940  
Qy 6058 AGCTCAGTGGAAACGAAACCTACGTTAAGGATTTTGGTCATGAGATTTATCAAAAGGA 6117  
Db 4941 AGCTCAGTGGAAACGAAACCTACGTTAAGGATTTTGGTCATGAGATTTATCAAAAGGA 5000  
Qy 6118 TCTTCACTAGATCTCTTTTAAATTAAGTTTAAATCAATCTAAAGTATATATG 6177  
Db 5001 TCTTCACTAGATCTCTTTTAAATTAAGTTTAAATCAATCTAAAGTATATATG 5060  
Qy 6178 AGTAACTTTGGTCTGACAGTTACCAATCTTAATCAGTGAGGACCTATCTCAGCGATCT 6237  
Db 5061 AGTAACTTTGGTCTGACAGTTACCAATCTTAATCAGTGAGGACCTATCTCAGCGATCT 5120  
Qy 6238 GTCTATTTCTGTTTATCCATAGTTGCTTGAATCTCCCGTCTGTGTAGATACTACGATACGG 6297  
Db 5121 GTCTATTTCTGTTTATCCATAGTTGCTTGAATCTCCCGTCTGTGTAGATACTACGATACGG 5180  
Qy 6298 AGGGCTTACATCTGGCCCCAGTGTGCAATGATACCGGAGACCACTCAGCCGGCTC 6357  
Db 5181 AGGGCTTACATCTGGCCCCAGTGTGCAATGATACCGGAGACCACTCAGCCGGCTC 5240  
Qy 6358 CAGATTTATCAGCAATAAACCCAGCAGCGGAGGCGGAGGAGAGTGTCTCTGCAA 6417  
Db 5241 CAGATTTATCAGCAATAAACCCAGCAGCGGAGGCGGAGGAGAGTGTCTCTGCAA 5300  
Qy 6418 CTTTATCCGCTTCACTCCAGTCTAATTAATTTGTTCGGGAAGCTAGAGTAGTTCCG 6477  
Db 5301 CTTTATCCGCTTCACTCCAGTCTAATTAATTTGTTCGGGAAGCTAGAGTAGTTCCG 5360  
Qy 6478 CAGTTAATAGTTTGGCAACGTTGTGCAATGCTTACAGGC-TCGTGTGTACCGCTCGT 6536  
Db 5361 CAGTTAATAGTTTGGCAACGTTGTGCAATGCTTACAGGCATCGTGTGTACCGCTCGT 5420  
Qy 6537 CGTTTGGTATGGCTTCACTCAGCTCCGTTCCCAACGATCAAGGCGAGTTACATGCC 6596  
Db 5421 CGTTTGGTATGGCTTCACTCAGCTCCGTTCCCAACGATCAAGGCGAGTTACATGCC 5480  
Qy 6597 CCATGTTTGCAAAAAGCGTTAGCTCTTCTCGGTCTCCGATCGTTGTTCAGAAAGT 6656  
Db 5481 CCATGTTTGCAAAAAGCGTTAGCTCTTCTCGGTCTCCGATCGTTGTTCAGAAAGT 5540  
Qy 6657 TGGCCGAGTGTATACATCATGTTATGGCAGCACTGCATTAATTTCTTCTTCTGTCTGC 6716  
Db 5541 TGGCCGAGTGTATACATCATGTTATGGCAGCACTGCATTAATTTCTTCTTCTGTCTGC 5600  
Qy 6717 CATCCGTAAAGTGTCTTTCTGTGACTGTGTGAGTACTCAACCAAGTCAATTCAGAAATAGT 6776  
Db 5601 CATCCGTAAAGTGTCTTTCTGTGACTGTGTGAGTACTCAACCAAGTCAATTCAGAAATAGT 5660  
Qy 6777 GTATCGGCGACCGAGTTGCTCTTCCCGGCTCAATACGGGATTAATACCGGCCACATA 6836  
Db 5661 GTATCGGCGACCGAGTTGCTCTTCCCGGCTCAACCGGATTAATACCGGCCACATA 5720  
Qy 6837 GCAGAACTTTAAAGTGTCTCATCATGTTGMAAAGCTTCTTCGGGGGGAAGAACTCTCAGGA 6896  
Db 5721 GCAGAACTTTAAAGTGTCTCATCATGTTGMAAAGCTTCTTCGGGGGGAAGAACTCTCAGGA 5780



QY 6897 TCTTACCGCTGTGAGATCCAGTTCGATGATTAACCACTCGTGCCACCAACTGATCTTCAG 6956  
Db |||||||  
QY 5781 TCTTACCGCTGTGAGATCCAGTTCGATGATTAACCACTCGTGCCACCAACTGATCTTCAG 5840  
Db |||||||  
QY 6957 CATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCA 7016  
Db |||||||  
QY 5841 CATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCA 5900  
Db |||||||  
QY 7017 AAAAGGGAATTAAGGCGGACAGGAAATGTTGAATACTCATCTCTCTCTCTCTCTCAATATT 7076  
Db |||||||  
QY 5901 AAAAGGGAATTAAGGCGGACAGGAAATGTTGAATACTCATCTCTCTCTCTCTCAATATT 5960  
Db |||||||  
QY 7077 ATTGAAGCATTTATCAGGCTTATTGTTCTCATGAGCGGATACATATTTGAATGTTATTAGA 7136  
Db |||||||  
QY 5961 ATTGAAGCATTTATCAGGCTTATTGTTCTCATGAGCGGATACATATTTGAATGTTATTAGA 6020  
Db |||||||  
QY 7137 AAAATAAACAATAGGCGTTCGGGACATTTCCCGAAAAAGTGCACCTGACGCTTAAG 7196  
Db |||||||  
QY 6021 AAAATAAACAATAGGCGTTCGGGACATTTCCCGAAAAAGTGCACCTGACGCTTAAG 6080  
Db |||||||  
QY 7197 AAACCATTTATTATCATGACATTAACCTATAAAATAGGCGTATCACGAGGCGCTTTTCGTC 7256  
Db |||||||  
QY 6081 AAACCATTTATTATCATGACATTAACCTATAAAATAGGCGTATCACGAGGCGCTTTTCGTC 6140  
QY 7257 T 7257  
Db 6141 T 6141

## RESULT 10

US-08-786-531B-3  
; Sequence 3, Application US/08786531B  
; Patent No. 6541197  
; GENERAL INFORMATION:  
; APPLICANT: Link, Charles J.  
; APPLICANT: Levy, John P.  
; APPLICANT: Wang, Suming  
; APPLICANT: Seragina, Tatiana  
; TITLE OF INVENTION: Vehicles for Stable Transfer of Green  
; TITLE OF INVENTION: Fluorescent Protein Gene and Methods of Use for Same  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Zarley, McKee, Thonite, Voorhees & Sease  
; STREET: 801 Grand Suite 3200  
; CITY: Des Moines  
; STATE: Iowa  
; COUNTRY: United States  
; ZIP: 50309  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/786.531B  
; FILING DATE: 21-JAN-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/010371  
; FILING DATE: 22-JAN-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Nebel, Heidi S.  
; REGISTRATION NUMBER: 37,719  
; REFERENCE/DOCKET NUMBER: hgttri  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 515-288-3667  
; TELEFAX: 515-288-1338  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6620 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

; MOLECULE TYPE: other nucleic acid  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
US-08-786-531B-3

Query Match 39.9%; Score 3050.2; DB 4; Length 6620;  
Best Local Similarity 70.4%; Pred. No. 0;  
Matches 4835; Conservative 0; Mismatches 1448; Indels 586; Gaps 33;

QY 549 TTTTGAAGAGACCCACACCTGTAGGTTTGGCAAGCTAGCTTAAAGTAACGCGCATTTTCAAGG 608  
Db |||||||  
QY 174 TTTTGAAGAGACCCACACCGTAGG--TGGCAAGCTAGCTTAAAGTAACGCGCATTTTCAAGG 231  
Db |||||||  
QY 609 CATGG-AAAATACATAACTGAGAAATAGAGAAAGTTTCAGATCAAGGTTAGGAAACAGAGAC 667  
Db |||||||  
QY 232 CATGGAAAATACATAACTGAGAAATAGAGAAAGTTTCAGATCAAGGTTAGGAAACAGAGAAAC 291  
Db |||||||  
QY 668 AGCAGATATGGGCAACACAGGATATCTGTGTAAAGCAGTTTCTGCCCC--GCTCAGGGCC 726  
Db |||||||  
QY 292 AGCTGAATA----CCAAACAGGATATCTGTGTAAAGCGGTTCTGCCCCGCTCAGGGCC 347  
Db |||||||  
QY 727 AAGAACAGTTTGAACACAGGAGAAAT-ATGGGCCAAACACAGGATATCTGTGTAAAGCAGTTTCT 785  
Db |||||||  
QY 348 AAGAACAGATGAGACAGCTGAGTGTATGGGCCAAACAGGATATCTGTGTAAAGCAGTTTCT 407  
Db |||||||  
QY 786 GCCCCGGCTCAGGCGCCAAAGAACAGATGTTCCCAAGATGCGGTCCCGCTCCAGCAGTTTC 845  
Db |||||||  
QY 408 GCCCCGGCTCAGGCGCCAAAGAACAGATGTTCCCAAGATGCGGTCCAGCCCTCAGCAGTTTC 467  
Db |||||||  
QY 846 TAGAAGACCATCAGATGTTTCCAGGGTGCCCAAGGACCTG-AAATGACCTGTGCGCTTA 904  
Db |||||||  
QY 468 TAGTGAATCATCAGATGTTTCCAGGGTGCCCAAGGACCTGAAATGACCTGTGACCTTA 527  
Db |||||||  
QY 905 TTTGAACCTAACCAATCAGTTTCGCTTCTCGCTTCTGTTTCGCGCGCTTCTGCTCCCGAGCT 964  
Db |||||||  
QY 528 TTTGAACCTAACCAATCAGTTTCGCTTCTCGCTTCTGTTTCGCGCGCTTCTGCTCCCGAGCT 587  
Db |||||||  
QY 965 CAATAAAGAGACCCACAAACCCCTCACTCGGCGGCCAGTCTCCGATAGACTGGGTCCGCC 1024  
Db |||||||  
QY 588 CAATAAAGAGACCCACAAACCCCTCACTCGGCGGCCAGTCTTCCGATAGACTGGGTCCGCC 647  
Db |||||||  
QY 1025 CGGTACCCGTATTTCCCAATAAAGCCTTCTGCTTTGTCATCCGAATCGTGAGACTCGCTG 1084  
Db |||||||  
QY 648 CGGTACCCGTATTTCCCAATAAAGCCTTCTGCTTTGTCATCCGAATCGTGAGACTCGCTG 707  
Db |||||||  
QY 1085 ATCCTTGGGAGGCTCCTCAGATTTGATTTGACTCCCACTCGGCGGCTTCTTCAATTGGA 1144  
Db |||||||  
QY 708 TTTCTTGGGAGGCTCCTCCTGAGTGATTTGACTACCCACGACGGGGGCTTTTTCATTGGG 767  
Db |||||||  
QY 1145 GGTTCACCGAGATTTGGAGACCCCTCGCCAGGAGACCAACCGACCCCGCCCGGAGGTA 1204  
Db |||||||  
QY 768 GGCTCGTCCGGATTTGGAGACCCCTCGCCAGGAGACCAACGA-CCACACCGGAGGTA 826  
Db |||||||  
QY 1205 AGTGGCCAGCAACTTATCTGTGTCGTCCGATTTGTCTAGTGTCTATGACTGATTTTATG 1264  
Db |||||||  
QY 827 AGCTGGCCAGCAACTTATCTGTGTCGTCCGATTTGTCTAGTGTCTATGTTTGTATG 886  
Db |||||||  
QY 1265 CGCTGCGTGGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGGACCGGTGTTGGAAT 1324  
Db |||||||  
QY 887 CGCTGCGTCTGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGGACCGGTGTTGGAAT 946  
Db |||||||  
QY 1325 GACGAGTTCCGAAACACCCCGCCGCAACCTCGGGAGAGAGTCCACGGAGCTTCGCGGGCGGT 1384  
Db |||||||  
QY 947 GACGAGTTCTGACACACCCCGCCGCAACCTCGGGAGAGAGTCCACGGAGCTTCGCGGGCGGT 1006  
Db |||||||  
QY 1385 TTTTGGCCCGGACCTGAGTCTTAAATCCCGATCGGTTTATAGGACTCTTTGTGACACCC 1444  
Db |||||||  
QY 1007 TTTTGGCCCGGACCTGAGGAGGAGTTCGATGTGGAATCCGACCCCGTCA----- 1057  
Db |||||||  
QY 1445 CTTAGAGGAGGATATGTTCTGTGTAGGAGAGAGACCTTAAACAGTTTCCCGCTCC 1504  
Db |||||||  
QY 1058 -----GGATATGTTGTTCTGTGTAGGAGAGAGACCTTAAACAGTTTCCCGCTCC 1107  
Db |||||||





3262	Db	GGTCTTGTGCAATCAGGATGATCTCGACGAAGAGCATCAGGG--GCTCGCGCCAGCCGGAAC	3319
3580	Qy	CCTCAGGAAGCCCTGTCAATGAACTCGAGAAAGATAGATGGCGGAGGCGCTACAGTGAAG	3639
3320	Db	TGTTGCGCCAGGCTCAAGCGCGCATGCCCGAGCGGAGGATCTCGTCTGTGAACCATGGCG	3379
3640	Qy	ATTGGGATGAAGGCGAGCGCGGAGGGGCAAGG--GGCAGATGGCCTTTTACAGGGGT	3697
3380	Db	ATGCCTGCTTTGCCGAATATCATGGTGAAATATGGCCGCTTTTCTGGAATTCATCGACTGTG	3439
3698	Qy	TCAGTACAGCCACCAGAGACACCTTACGACGCCCTTCACATGCGAGGCCCTGCCCCCTCGCT	3757
3440	Db	GCCGGCTGGGTGGCGGACCGCTATCAGAGACATAGCGTTGGCTACCCCGTGATATTGCTG	3499
3758	Qy	AA-----CTCGACGCGCGCGCGGATCCGGATTTAGTCCAATTTGTTTAAAGACAGGATATCAG	3813
3500	Db	AAGAGCTTTGGCGGCGAATGGCTGACCGCTTCTCTGTCCTTAAGGTATCGCCGCTCCCG	3559
3814	Qy	TGGTCCAGGCTCTAGTTTGTGACTCAACAATATCACAGCTGAAGCCCTATAGAGTACGAGC	3873
3560	Db	ATTGCGACGCGCATCGCCTTCTATCGCCTTCTTTCACGAGTTCCTCTGAGCGGACTCTGGG	3619
3874	Qy	CATAGATAAATAAAGATTTTATTTAGTCTCCAGAAATAAGGGGGGAATCAAGAGCCCA	3933
3620	Db	GTTTCGATAAATAAAGATTTTATTTAGTCTCCAGAAATAAGGGGGGAATGAAAGACCCCA	3679
3934	Qy	CCTGTAGCTTTGGCAAGCTAGCTTAAAGTAACGCCATTTTCGCAAGGCATGG--AAATATACAT	3992
3680	Db	CCTGTAGCTTTGGCAAGCTAGCTTAAAGTAACGCCATTTTCGCAAGGCATGAAATAATACAT	3739
3993	Qy	AATCGAATAGAGAAGTTCCAGATCAAGTTTAGGAACAGA--GAGACAGCAGAGATATGGGC	4051
3740	Db	AACGTGAATAGAGAAGTTCCAGATCAAGTTCAGGAACAGATGGAACAGCTGAAATATGGGC	3799
4052	Qy	CAAACAGGATATCTGTGGTAGCAGTTTCCTGCCCC--GCTCAGGGCCCAAGAACAGTTGGAA	4110
3800	Db	CAAACAGGATATCTGTGGTAGCAGTTTCCTGCCCCGCTCAGGGCCCAAGAACAGATGGAA	3859
4111	Qy	CAGGAGAAATATGGGCCAAACAGGATATCTGTGGTAGCAGTTCTCTGCCCGGCTCAGGGC	4170
3860	Db	CAGCTGAATATGGGCCAAACAGGATATCTGTGGTAGCAGTTCTCTGCCCGGCTCAGGGC	3919
4171	Qy	CAAGAACAGATGTTCCCGATGCGGTCCCGCCTTCAGCAGTTTCTTAGAGAACCATCAGA	4230
3920	Db	CAAGAACAGATGTTCCCGATGCGGTCCAGCCTCAGCAGTTTCTTAGAGAACCATCAGA	3979
4231	Qy	TGTTTTCAGGGTGCCCAAGACCTGAATGACCCCTGTGCTTATTTTGAATCAACCATC	4290
3980	Db	TGTTTTCAGGGTGCCCAAGACCTGAATGACCCCTGTGCTTATTTTGAATCAACCAATC	4039
4291	Qy	AGTTTCGGTCTTCGCTTCTGTTTCGGCGCTTCTGCTCCCGAGCTCAATAAAGAGCCAC	4350
4040	Db	AGTTTCGGTCTTCGCTTCTGTTTCGGCGCTTCTGCTCCCGAGCTCAATAAAGAGCCAC	4099
4351	Qy	AACCCCTCACTCGGCGCCAGTCTCTCCGATAGACTCGTCCGCCGGGTACCCGCTGTTC	4410
4100	Db	AACCCCTCACTCGGCGCCAGTCTCTCCGATAGACTCGTCCGCCGGGTACCCGCTGTATC	4159
4411	Qy	CAATAAACCCCTCTTGAGATTGCATCCGACTCGTGGTCTCGCTGTTCTCTGGGAGGGTCT-	4469
4160	Db	CAATAAACCCCTCTTGAGATTGCATCCGACTCGTGGTCTCGCTGTTCTCTGGGAGGGTCTC	4219
4470	Qy	CTCTGAGTGATTGACTACCGTTCAGCGGGTCTTTTCAGTTTCTCCACCTTACACAGTCT	4529
4220	Db	CTCTGAGTGATTGACTACCGTTCAGCGGGGTCTTTTCAGTTTGGGGGTCTCGTCCGGGATCG	4279
4530	Qy	CACATAATCTCTGATGTGCGCAGGAGCTCCGTCAGCCCGGTTTTTTTGTATTATAAATAA	4589
4280	Db	GGAGACCCCTCCCGCAGGAGCAACCGACCCACCGGAGGTAAAGCTGG-----4328	
4590	Qy	TGCAAGAACAGTGTTCCTCTCAAGCCAGACTACATATCTGACTCTCGGCTTTTATAAAGAA	4649
4329	Db	-----CTGCTCTCGCGGTTTCGGTGATGA4352	

Qy	4650	TGTTGAAGGGCTCTGTGGACTATCTGTGCCACACGACTTTTAAAGATTTTATGAGCTCTCTGG	4709
Db	4353	CGGTGAAAACCTCTG-ACACATCAGCTCCGGAGACGGTACACAGCTTGTCTCTGAAGCGG	4411
Qy	4710	ATGAGGATTTAGTCAATCTATCCTCGTCTATTTTCTGCTTCTCGATTTTAAATTT	4769
Db	4412	ATGCCGGAGCAGACAAGCCCGTCAAGGCGCTCAGCGGTGTGGCG-----	4459
Qy	4770	CTAGTTTGCACCTCCCTTCTTGAGAGCAGCGCATGTCAGAGTAGTTAACTACTCTGAGGC	4829
Db	4460	-----	4459
Qy	4830	AGGCTTCTGTGAAAAGGTTCCCTGGGCTCAGTGTGAGATTTTCCCATAAAAAGGGGTCT	4889
Db	4460	-----	4459
Qy	4890	GCCCTGTGTACAGACAGATCGGAATCTAGAGTGCATCTCAGAGTCCCGCGGTCCGG	4949
Db	4460	GGCTCTGATCTCAGGGCATCTTTTCCCTAGAGATCTCTACCGCGACGCACTCGTGCCGG	4467
Qy	4950	GGCTCTGATCTCAGGGCATCTTTTCCCTAGAGATCTCTACCGCGACGCACTCGTGCCGG	5009
Db	4468	GGCGCACCATGACCCAGTC-----ACGTAGCATAGCGG	4502
Qy	5010	GTACCGAGCTCGAATTGTTAATCATGGTCATAGTCTGTTCCTGTGTAAATTTTATCCG	5069
Db	4503	AGTGTATACTGGCTTAACTATGCGGCATCAGACGAGATGTTACTGAGA-----	4550
Qy	5070	CTCAAAITTCACACAACATACAGCGGAGAGCATAAAGTGTAAAGCTGGGTGCTTAA	5129
Db	4551	-----	4550
Qy	5130	TGAGTGAGTTAACTCACATTTAATTGGTGTGGCTCACTGCGCTTTTCCAGTCGGGAAAC	5189
Db	4551	-----	4550
Qy	5190	CTGTCTGCGCAGCTGCATTAATGAATCGGCCAACGGCGGGAGAGCGGTTTGCGTATT	5249
Db	4551	---GTGCACCATATGGGTGTGAATAACCGCACAGATGCTAAGGAGAAATACCGCATC	4607
Qy	5250	GGCGCTCTTTCGGTCTCTCGCTCACTGTACTCGCTGGCTCGGTCTGTTTGGCTGCGGCGA	5309
Db	4608	AGCGCTCTTTCGGCTTCTCGCTCACTGTACTCGCTCGCTCGGTCTGCGTTCGGCTGCGGCGA	4667
Qy	5310	GCGGTATCAGTCACTCAAAAGGCGGTAAATACGGTTATCCACAGAAATCAGGGGATACGCA	5369
Db	4668	GCGGTATCAGTCACTCAAAAGGCGGTAAATACGGTTATCCACAGAAATCAGGGGATACGCA	4727
Qy	5370	GGAAGAAATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGGTTG	5429
Db	4728	GGAAGAAATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGGTTG	4787
Qy	5430	CTGCGTTTTCATAGGCTCGCGCCCTTGACGAGCATCACAAAATCGACGCTCAAGT	5489
Db	4788	CTGCGTTTTCATAGGCTCGCGCCCTTGACGAGCATCACAAAATCGACGCTCAAGT	4847
Qy	5490	CAGAGTGGGAAACCCGACAGAGACTATAAGATACAGGCGTTTCCCGCTGGAAGCTCC	5549
Db	4848	CAGAGTGGGAAACCCGACAGAGACTATAAGATACAGGCGTTTCCCGCTGGAAGCTCC	4907
Qy	5550	CTGTGCGCTCTCTGTTCGACCCCTTCACCGGATACCTGTCCGCTTTCTCCCT	5609
Db	4908	CTGTGCGCTCTCTGTTCGACCCCTTCACCGGATACCTGTCCGCTTTCTCCCT	4967
Qy	5610	TCGGGAAGCGTGGCGCTTTCTCATAGTCACTCGCTGTAGGTATCTCAGTTTGGTGTAGGTC	5669
Db	4968	TCGGGAAGCGTGGCGCTTTCTCATAGTCACTCGCTGTAGGTATCTCAGTTTGGTGTAGGTC	5027
Qy	5670	GTTCGCTCCAAGCTGGGCTGTGTGCAGAAACCCCGCTTCAGGCCAGCGCTCGGCGTTA	5729
Db	5028	GTTTCGCTCCAAGCTGGGCTGTGTGCAGAAACCCCGCTTCAGGCCAGCGCTCGGCGTTA	5087



[illegible]

1229	ATCACCCAGGTTAAGATCAAGGCTCTTTTTCACCTGGGCCCGCATGGACACCCAGGACCCAGGTC	1289
1863	CCCTACATCGTGACCTGGGAAGCCTTCGCTTTTTCAGCCCGCTCCCTCGGTGCAAGCCCTTT	1922
1289	CCCTACATCGTGACCTGGGAAGCCTTCGCTTTTTCAGCCCGCTCCCTCGGTGCAAGCCCTTT	1348
1923	GTACACCCATAAGCCTCCGCCTCTCTTCTTCCATTCGCCCCCGTCTCTCCCCCTTGAACCT	1982
1349	GTACACCCATAAGCCTCCGCCTCTCTTCTTCCATTCGCCCCCGTCTCTCCCCCTTGAACCT	1408
1983	CCTCGTTGACCCCGCTCGATCTCTCGCTTTTATCCAGCCCTCACTCTTCTCTAGGGCC	2042
1409	CCTCGTTGACCCCGCTCGATCTCTCGCTTTTATCCAGCCCTCACTCTTCTCTAGGGCC	1468
2043	CCCATATGGCCATATGAGATCTTATATGGGGACCCCGCCCTTGTTAAACTTCCCTGAC	2102
1469	-----GGAAATTGCGGCGCGCTAGCTAGTCGACTCGC	1499
2103	CCTGACATGACAAGAGTTACTTAACAGACCCCTCTCTCCAAGCTCACTTACAGGCTTCTACT	2162
1500	TGTGGAATGTGTGTCAGTTAGGGGTGGAAAGTCCCAGGCTCCCCAGCAGGC-----	1552
2163	TAGTCCAGCAGCAAGTCTGGAGACCTCTGGGGGAGCCTACCAAGAACAACTGGACGGAC	2222
1553	-----AGAAAGTATGCAAA-----	1565
2223	CGGTGGTACCTACCCCTTACCGAGTGGCGACACAGTGTGGTCCGCGACACCAAGACTA	2282
1566	-----GCATGCATCTCAATTAGTCAGAACACAGGTGTGGAAAGTCCCAGGCTCCCCA	1618
2283	AGAACTAGAACCTCGCTGGAAAGGACCTTACACAGCTCTGTGACCAACCCCCACCGCCC	2342
1619	GCAGCAGAAGTATGCAAGCATGCATCTCAATTAGTC-----AGCAACCATAGTCCCGC	1673
2343	TCAAAGTAGCGGATCGAGCTTGGATACACGGCGGCCACAGTGAAGGTGCGGACCCCG	2402
1674	CCCTAACTCCGCCCATCCCGCCCTAACTCCGCCGAGTTCGCGCCCATTTCTCCGCCCATG	1733
2403	GGGTGGACCATCCTCTAGAGCTGCATGGGATGGAGCTGTATCATCTCTCTTGTGTAGC	2462
1734	GCTGACTAAATTTTTTATTATGACAGAGCCGAGGCGCCTCGGCTCTTGAGCTATTC	1793
2463	AACAGCTACAGGTGTCCACTCCGACATCCAGCTGACCCAGAGCCCAAGCAGCCTGAGCGC	2522
1794	AGAAG---TAGTGAGGAGGCTTTTGGAGGCTTAGGCTTTTGGCAAAAAGCTCGAAGAT	1849
2523	CAGCGTGGGTGAAGAAGTGACATCAGCTGTAAAGCCAGTCAGGATGTGGGTACTTCTGT	2582
1850	CAATTCCGCTATGATCAAGA-----GACAGGATGAGGATCGTTTCGC	1891
2583	AGCTTGTGACAGCAGAGCCAGGTAAAGGCTCCAAAGCTGCTGATCTACTGGACATCCAC	2642
1892	ATGATTGAAACAAGATGATTTGACAGCAGGTTCTCCGCGCGCTTGGGTGGAGAGGCTATTTC	1951
2643	CCGCGCACACTGGTGTGCGCAAGACAGATTACAGCGGTAGCGGTAGCGGTACCGACTTCACCTT	2702
1952	GGCTATGACTGG-----	1963
2703	CACCATCAGCAGGCTCCAGCCAGAGGACATCGGCACTACTACTGCAGCAATATAGCCT	2762
1964	---GCACAAACAGACAATCTGGCTCTCTGTATGCCGCGCTGTTCGCGCTGTGAGCGAGGGG	2020
2763	CTATCGGTCTGCGCCACGGGACCAAGGTGGAAATCAACAGAGGTGGCTCAGGATCGGG	2822
2021	CGCCCGGTCTTTTGTGCAAGACGACCTGTCCGCTGCGCTGAAATGAACTGCAGGACGAG	2080
2823	TGGATCCGGCTCTGGTGGCTCAGGATCGGAGTCCAACTGGTGGAGAGCGGTGGAGGTGT	2882
2081	GCAGCGCGGCTATCGTGGCT-----GGCCACGACGGCGTTCCT	2119
2883	TGTGCAACCTGGCGGTCCTCGCGCTGCTGCTCCGATCTGGCTTCGATTCACACAC	2942
2120	TGCGCAGCTGTGCTGCACGTGTGTCATGAAAGCGGGAAGGACCTGGCTGTAT-----	2171

QY 2943 ATATTGGATGAGTGGGTGAGACAGGACCTCGAAAGGCTCTGAGTGGATTGGAGAAAT 3002  
 Db |||||  
 2172 ---TGGCGAAATGCGGGGAGGATCTCTGTCTATCTACCTTGTCTCTGCGCGAGAAAG 2228  
 QY TCATCCAGATAGACGATACGATTAACTATGCGCGCTCTCTAAAGGATAGATTTACAATATC 3062  
 Db |||||  
 2229 TATCCATCATGCTGATGCA-----ATGCGGGGCTGCATACGCTTGAT----- 2272  
 QY 3063 GCGAGACACGCCAAGAACACATGTTCTGTGCAATGAGACAGCTGAGACCCGAGACAC 3122  
 Db -----CCGGTACTCGCCCAATCGACCAACCAAGCAAAACATCGCATCGAGC 2318  
 QY 3123 CGGGGTCTATTTTGTGCAAGCTTTACTTGGCTTCCCTGTGTTGCTTTATGCGGCA 3182  
 Db |||||  
 2319 GAGCAGTACTCGGATGAGCGGCTTGTGATCAGATGATCTGAGCAAGAGCATC 2378  
 QY 3183 AGGGAACCCGGTACCGTCTCAGTCTAAGCCCAACGACGCGAGCGCCGCGACAC 3242  
 Db |||||  
 2379 AGGGGCTCGGCG-----AGCGGAACCTGTTGCGCAGGCTCAAGCGGCGC 2422  
 QY 3243 AACACGGGCGCCACCATCGCTCGAGCCCTGTGCTGCGCCGCGCAGAGCGGCTCGGCC 3302  
 Db |||||  
 2423 ATGCCGACGCGGAGGATCTCGTGTGACCCATGCG----- 2458  
 QY 3303 AGCGGGGGGCGCAGTGCACACAGAGGGGCTGGAATTCGCGCTGGATCCCAAACTCTG 3362  
 Db |||||  
 2459 -----GATGCTGTGCTGCC 2472  
 QY 3363 CTACTGCTGGATGGAATCTCTTCACTATGATGTTGATCTCACTGCTGCTGCTGAG 3422  
 Db |||||  
 2473 GAATATCATGGTGGAAATGCGGCTTTTCTGGAATCATCGACTGTG----- 2519  
 QY 3423 AGTGAAGTTCAGCAGAGCGCAGAGCGCCCGCGGTACGAGCGGCGCAGAACCGACTCTA 3482  
 Db |||||  
 2520 -----GCCGGCTGGGTGTGGCGACCGCTA 2544  
 QY 3483 TAACGAGCTCAATCTAGGACGAAGAGAGGAGTACGATGTTTGGCAAGAGACGTGGCG 3542  
 Db |||||  
 2545 TCAGG----- 2549  
 QY 3543 GGACCTGAGATGGGGGAAAGCCGAGAGAGAAACCTCAGAGAGGCTGTACAATGA 3602  
 Db ----- 2549  
 QY 3603 ACTGCAGAAAGATAAGATCGCGAGGCGCTACAGTGAATGGGATGAAGGCGAGGCGCG 3662  
 Db |||||  
 2550 -----ACATAGCGTTGGCTACCGGTGATATTG-----CT 2578  
 QY 3663 GAGGSCAAGGGGCGACGATGGCTTTTACCAGGCTCTCAGTACAGCCACCAAGGACACCTA 3722  
 Db |||||  
 2579 GAAGACTTGGCGCGAAATGGGCT-----GACCGCTT 2610  
 QY 3723 CGACGCTTTCATGACAGGCGCTGCGCCCTCGCTAACTCGACGCGCGCGGATCCGGA 3782  
 Db |||||  
 2611 CCTCGTCTTTACGTTATCGCGCTCCGATTCGAGCGCATCGCTTCTATCGCTTCT 2670  
 QY 3783 TTAGTCCAAATTTGTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTGTGACTCAAA 3842  
 Db |||||  
 2671 TGACGAGTCTTCTTGAGCGGACTCTGGGGTTCGTGCGAAGCTTG----- 2716  
 QY 3843 TATCACCAGCTGAGCCTATAGGTACGAGCCATAGATAAATAAAGATTTTATTAGT 3902  
 Db -----GGCCATCGATAAATAAAGATTTTATTAGT 2749  
 QY 3903 CTCAGAAAAGGGGGGAATGAAGACCCACCTGTAGGTTTGGCAAGCTAGCTTAAGTA 3962  
 Db |||||  
 2750 CTCAGAAAAGGGGGGAATGAAGACCCCACTGTAGGTTTGGCAAGCTAGCTTAAGTA 2809  
 QY 3963 AGCCATTTTGCNAGGCTAGG-AAAATACATACTGAGAAATAGAGAGTTTCAGTCAAGG 4021  
 Db |||||  
 2810 AGCCATTTTGCNAGGCTAGGAAAAATACATACTGAGAAATAGAGAGTTTCAGTCAAGG 2869

QY 4022 TTAGGAAACAGA-GAGACACAGAAATATGGGCCAAACAGGATATCTGTGGTAAAGCATTTCC 4080  
 Db |||||  
 2870 TCAGGAAACAGATGGAACAGCTGAATATATGGGCCAAACAGGATATCTGTGGTAAAGCATTTCC 2929  
 QY 4081 TGCCCC-GCTCAGGGCCAAAGAACAGTTGGAAACAGGAGAAATATGGGCCAAACAGGATATCT 4139  
 Db |||||  
 2930 TGCCCCGGCTCAGGGCCAAAGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCT 2989  
 QY 4140 GTGGTAAGCAGTCTCTGCCCGCTCAGGGCCAAAGAACAGATGTTTCCCGATGGCTGCC 4199  
 Db |||||  
 2990 GTGGTAAGCAGTCTCTGCCCGCTCAGGGCCAAAGAACAGATGTTTCCCGATGGCTGCC 3049  
 QY 4200 CGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCCCAAGAACCTGAAA 4259  
 Db |||||  
 3050 AGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCCCAAGAACCTGAAA 3109  
 QY 4260 TGAACCTGTGCTTATTTGAACTAAACAAATCAGTTGCTTCTCGCTTCTGTTCGCGCGCT 4319  
 Db |||||  
 3110 TGACCTGTGCTTATTTGAACTAAACAAATCAGTTGCTTCTCGCTTCTGTTCGCGCGCT 3169  
 QY 4320 TCTGCTCCCGAGCTCAATAAAGAGCCCAACACCCCTCAGTTCGGCGCGCCAGTCTCCG 4379  
 Db |||||  
 3170 TCTGCTCCCGAGCTCAATAAAGAGCCCAACACCCCTCAGTTCGGCGCGCCAGTCTCCG 3229  
 QY 4380 ATAGACTCGCTCGCCCGGTACCCGCTGTTCTCAATAAACCCCTCTTTCAGTTGCTCATCCGAC 4439  
 Db |||||  
 3230 ATTGACTGAGTCGCGCGGTACCCGCTGTTATCAATAAACCCTCTTTCAGTTGCTCATCCGAC 3289  
 QY 4440 TCGTGTCTCGTGTCTCTTGGGAGGCTCTCTCTGAGTGAATGACTACCGTTCAGCGGG 4498  
 Db |||||  
 3290 TTGTGTCTCGTGTCTCTTGGGAGGCTCTCTCTGAGTGAATGACTACCGTTCAGCGGG 3349  
 QY 4499 GTCCTTTCAGTTTCTCCACCTACACAGCTCAGTCACTAACATTCCTGATGTCGCGAGGAC 4558  
 Db |||||  
 3350 GGTCTTTCATTTGGGGGCTCGTTCGGGATCGGAGACCCCTGCCAGGAGCACCGACCC 3409  
 QY 4559 TCCGTGACCGCGTTTGTGTTTATAATAAATGCAAGAACAGTGTTCCTTTCAGCCAGA 4618  
 Db |||||  
 3410 ACCACCGGAGTGAAGCTG----- 3429  
 QY 4619 CTACATCTGACTCTCGGCTTTTAAAGAAATGTTGAAGGGCTCTGTGGACTATCTGCCA 4678  
 Db |||||  
 3430 -----CTGCTCGCGGTTTTCGGTGATGACGGTGAAGAAACCTCTG-ACATATGACGCTC 3481  
 QY 4679 CACGACTTTTAAAGATTTTATGCTCTGATGAGGGATTTAGTCAATCTATCTCTCTC 4738  
 Db |||||  
 3482 CCGGAGACGGTCAACAGTTGCTGTGAAGCGGATGCGGAGAGCAGAACAGCCGCTCAGGCG 3541  
 QY 4739 TATTTGTGCTGCTTCTCGCTATTTTAAATTTTCTAGTTTGCATCTCCCTTCTGAGAGCAG 4798  
 Db |||||  
 3542 GGTTCAGCGGGTGTGGCG----- 3560  
 QY 4799 GCGATTCAGAGTGTGTTAATCTCTGAGGGCAGGCTTCTGTGAAGAGTTGCTTGGGCTC 4858  
 Db |||||  
 3561 ----- 3560  
 QY 4859 AGTGTGAGATTTTGCCATAAAGGGGCTCTGCCCTCTGTGTACACAGATCGGAATCTA 4918  
 Db |||||  
 3561 ----- 3560  
 QY 4919 GAGTGCATACCTCAGAGTCCCGGCTTCCGGGCTCTGATCTCAGGGCATCTTTGCTGCTAG 4978  
 Db |||||  
 3561 -----GGTGTGCGGGCGCAGCCATGA----- 3581  
 QY 4979 AGATCTCTACGCGGACGCACTCGTGGCGGGGTACCGAGTCTGGAATCTGTAATCATGTGTC 5038  
 Db |||||  
 3582 ----- 3581  
 QY 5039 ATAGCTGTTTCTGTGTGAATTTGTTATCCGCTCAAAATCCACACAACATACGAGCCGG 5098  
 Db |||||  
 3582 -----CCAGTTCAGTACGATAGCG 3602  
 QY 5099 AAGCATAAAGTGAAGGCTCGGGGTGCTTAATGATGAGTAACTCACAATTAATTAATTCGCTT 5158



Db 3603 GAGTGTATCTGGCTTAACATATGCGGCATCAGACGAGATTGTACTGAGAGT----- 3653  
Qy 5159 GCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTGTCGCCAGCTGCATTAATGAATCGG 5218  
Db 3654 -----GCACCATATGCGGTGTGAATACC 3677  
Qy 5219 CCAACGCGCGGGAGAGCGGTTTGGCTATTGGGGCTCTTCGCGTTCCTCGCTCACTGA 5278  
Db 3678 GCACAGATGCGTAAGGAGAAATACCGCATCAGGCGCTCTTCGCGTTCCTCGCTCACTGA 3737  
Qy 5279 CTCGCTGCGCTCGCTGCTGCTGCGGAGCGGTATCAGTCACTCAAAAGGCGGTAAAT 5338  
Db 3738 CTCGCTGCGCTCGCTGCTGCGGAGCGGTATCAGTCACTCAAAAGGCGGTAAAT 3797  
Qy 5339 ACGGTTATCCACAGAAATCAGGGGATAACCGCAGGAAAGAACATGTGAGCAAAAGGCCACGA 5398  
Db 3798 ACGGTTATCCACAGAAATCAGGGGATAACCGCAGGAAAGAACATGTGAGCAAAAGGCCACGA 3857  
Qy 5399 AAAGCCAGGAACCGTAAAGAGCGCGTTCGCTGCGGTTTTTCCATAGGCTCCGCGCCGCC 5458  
Db 3858 AAAGCCAGGAACCGTAAAGAGCGCGTTCGCTGCGGTTTTTCCATAGGCTCCGCGCCGCC 3917  
Qy 5459 TGACGAGCATCAAAATACAGCTCAAGTCAGAGTGGCGAAACCCGACAGACTATA 5518  
Db 3918 TGACGAGCATCAAAATACAGCTCAAGTCAGAGTGGCGAAACCCGACAGACTATA 3977  
Qy 5519 AAGATACAGAGGTTTTCCCGCTTGAAGCTCCCTCGTGCGCTCTCTGTTCCGACCCCTGCC 5578  
Db 3978 AAGATACAGAGGTTTTCCCGCTTGAAGCTCCCTCGTGCGCTCTCTGTTCCGACCCCTGCC 4037  
Qy 5579 GCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGCGAAGCGTGGCGCTTTCTCATAGCTC 5638  
Db 4038 GCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGCGAAGCGTGGCGCTTTCTCATAGCTC 4097  
Qy 5639 ACGCTGATAGGTATCTCAGTTCGCTGATAGTTCGCTCCAAAGCTGGGCTGTGTGCACGA 5698  
Db 4098 ACGCTGATAGGTATCTCAGTTCGCTGATAGTTCGCTCCAAAGCTGGGCTGTGTGCACGA 4157  
Qy 5699 ACCCGCGTTCAGCGCCGACGCTGCGCTTATCCGTTAACTATCGTTGAGTCCAAACC 5758  
Db 4158 ACCCGCGTTCAGCGCCGACGCTGCGCTTATCCGTTAACTATCGTTGAGTCCAAACC 4217  
Qy 5759 GGTAAAGACACACTTATCGCACTGGCAGCAGCCTGTTAAAGGATTAGCAGAGCGAG 5818  
Db 4218 GGTAAAGACACACTTATCGCACTGGCAGCAGCCTGTTAAAGGATTAGCAGAGCGAG 4277  
Qy 5819 GTATGATAGCGGTGTACAGAGTTCCTGAAGTGGTGGCTTAACTACGCGCTACACTAGAAG 5878  
Db 4278 GTATGATAGCGGTGTACAGAGTTCCTGAAGTGGTGGCTTAACTACGCGCTACACTAGAAG 4337  
Qy 5879 GACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTCGGAAAGAGTTGGTAG 5938  
Db 4338 GACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTCGGAAAGAGTTGGTAG 4397  
Qy 5939 CTCCTGATCCGGCAAAACAAACACGCTGCTGAGCGGTGTTTTTTTGTTCGAAGCAGCA 5998  
Db 4398 CTCCTGATCCGGCAAAACAAACACGCTGCTGAGCGGTGTTTTTTTGTTCGAAGCAGCA 4457  
Qy 5999 GATTACGCGCAGAAAAAAGGATCTCAAGAAAGATCTTTTGATCTTTTCTACGGGGTCTGA 6058  
Db 4458 GATTACGCGCAGAAAAAAGGATCTCAAGAAAGATCTTTTGATCTTTTCTACGGGGTCTGA 4517  
Qy 6059 CGCTCAGTGGAAACGAAACTCAGTTAAGGATTTTGTGTCATGAGATTATCAAAAAGGAT 6118  
Db 4518 CGCTCAGTGGAAACGAAACTCAGTTAAGGATTTTGTGTCATGAGATTATCAAAAAGGAT 4577  
Qy 6119 CTTACCTAGATCTTTTAAATTAAGTTTAAATCAATCTAAAGTATATGA 6178  
Db 4578 CTTACCTAGATCTTTTAAATTAAGTTTAAATCAATCTAAAGTATATGA 4637  
Qy 6179 GTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAAGGCACTTCTCAGCGGATCTG 6238

Db 4638 GTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAAGGCACTTCTCAGCGATCTG 4697  
Qy 6239 TCTATTTTGGTTCATCCATAGTTGCTGACTCCCGTCTGCTGATAGTAATACTACGATACGGGA 6298  
Db 4698 TCTATTTTGGTTCATCCATAGTTGCTGACTCCCGTCTGCTGATAGTAATACTACGATACGGGA 4757  
Qy 6299 GGGCTTACCATCTGGCCCGAGTGTCAATGATACCGGAGACCCACGCTCACGGCTCC 6358  
Db 4758 GGGCTTACCATCTGGCCCGAGTGTCAATGATACCGGAGACCCACGCTCACGGCTCC 4817  
Qy 6359 AGATTTATCAGCAATAAACACAGCCAGCGGAAGCGCGAGCGAGAGTGTCTCTGCAAC 6418  
Db 4818 AGATTTATCAGCAATAAACACAGCCAGCGGAAGCGCGAGAGTGTCTCTGCAAC 4877  
Qy 6419 TTTATCCGCTCCATCCAGTCTTATTAATTTGTCGGGGAAGCTAGAGTAAGTCTCGCC 6478  
Db 4878 TTTATCCGCTCCATCCAGTCTTATTAATTTGTCGGGGAAGCTAGAGTAAGTCTCGCC 4937  
Qy 6479 AGTTAATAGTTTGGGCAACGTTGTTGTCATTTGCTACAGGC-TCGTGTTGTCACGCTCGTC 6537  
Db 4938 AGTTAATAGTTTGGGCAACGTTGTTGTCATTTGCTGCAAGGCATCGTGTGTCACGCTCGTC 4997  
Qy 6538 GTTTGGTATGCTTCATTCAGTCCGCTCCAGATCAAGCGAGTTACATGATCCCC 6597  
Db 4998 GTTTGGTATGCTTCATTCAGTCCGCTCCAGATCAAGCGAGTTACATGATCCCC 5057  
Qy 6598 CATGTTGTGCAAAAAAGCGGTAGTCTCTTCGGTCTCCGATCGTTGTCAGAACTAAGTT 6657  
Db 5058 CATGTTGTGCAAAAAAGCGGTAGTCTCTTCGGTCTCCGATCGTTGTCAGAACTAAGTT 5117  
Qy 6658 GGGCGAGGTATCATCTCATGTTTGGCAGCACTGCATAAATTTCTCTTACTGTCATGCC 6717  
Db 5118 GGGCGAGGTATCATCTCATGTTTGGCAGCACTGCATAAATTTCTCTTACTGTCATGCC 5177  
Qy 6718 ATCCGTAAGATGCTTTTCTGCTGCTGCTGAGTACTCAACAGTCACTTCTGAGAAATAGTG 6777  
Db 5178 ATCCGTAAGATGCTTTTCTGCTGCTGCTGAGTACTCAACAGTCACTTCTGAGAAATAGTG 5237  
Qy 6778 TATCGCGGACCGAGTTGCTTCTGCGCGGTCAATACGGGATTAATACCGCGCCACATAG 6837  
Db 5238 TATCGCGGACCGAGTTGCTTCTGCGCGGTCAACACGGGATTAATACCGCGCCACATAG 5297  
Qy 6838 CAGAACTTTAAAGTGTCTCATTTGAAAAAGTTTCTTCGGGGCGGAAACTCTCAAGGAT 6897  
Db 5298 CAGAACTTTAAAGTGTCTCATTTGAAAAAGTTTCTTCGGGGCGGAAACTCTCAAGGAT 5357  
Qy 6898 CTTACCGCTGTTGAGATCCAGTTCGATGTAAACCACTCGTGCACCACTGATCTTCAGC 6957  
Db 5358 CTTACCGCTGTTGAGATCCAGTTCGATGTAAACCACTCGTGCACCACTGATCTTCAGC 5417  
Qy 6958 ATCTTTTACTTTTACACAGCGTTTCTGGGTGAGCAAAACAGGAAGGCAAAATGCGCAAA 7017  
Db 5418 ATCTTTTACTTTTACACAGCGTTTCTGGGTGAGCAAAACAGGAAGGCAAAATGCGCAAA 5477  
Qy 7018 AAAGGGAATTAAGCGGACACGGAATTTGAAATCTCATACTCTTCTCTTTTTCATATTA 7077  
Db 5478 AAAGGGAATTAAGCGGACACGGAATTTGAAATCTCATACTCTTCTCTTTTTCATATTA 5537  
Qy 7078 TTGAAGCATTTATCAGGGTTATTTGCTCATGAGCGGATACATATTTGAATGTTATTAAGAA 7137  
Db 5538 TTGAAGCATTTATCAGGGTTATTTGCTCATGAGCGGATACATATTTGAATGTTATTAAGAA 5597  
Qy 7138 AAATAAACAATAGGGTTCGCGCACATTTCCCGGAAAGTGCACCTGACGCTCTAAGA 7197  
Db 5598 AAATAAACAATAGGGTTCGCGCACATTTCCCGGAAAGTGCACCTGACGCTCTAAGA 5657  
Qy 7198 AACCATTTATCATGACATTAACCTATAAATAAGGCGTATCAGAGGCGCTTTCTGCT 7257  
Db 5658 AACCATTTATCATGACATTAACCTATAAATAAGGCGTATCAGAGGCGCTTTCTGCT 5717  
Qy 7258 CGCGGTTTCCGTTGATGACCGGTGAAAAACCTCTGACACATGACGCTCCCGGAGACGGTCA 7317  
Db 5718 TCAAGAATTCATACAGATCACCGGAAACTGTCTCCAAATGTGTCCCGCTCACAACCTCCC 5777



QY 7318 AGCTTCTGTAGCGGATGCGGGAGCAGCAAGCC 7354  
 Db 5778 AATTCCGGGGCTTCGCGCTCTTAGACCACCTCTACCC 5814

RESULT 12

US-09-935-194-1  
 ; Sequence 1, Application US/09935194  
 ; Patent No. 6635448  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bucciarelli  
 ; APPLICANT: Levenson  
 ; APPLICANT: Primiano  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INCREASING PROTEIN YIELD  
 ; FILE OF INVENTION: FROM A CELL CULTURE  
 ; FILE REFERENCE: 11202-3  
 ; CURRENT APPLICATION NUMBER: US/09/935,194  
 ; PRIOR FILING DATE: 2001-08-21  
 ; PRIOR APPLICATION NUMBER: 60/226,290  
 ; PRIOR FILING DATE: 2000-08-21  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 1  
 ; LENGTH: 7086  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: Expression  
 ; OTHER INFORMATION: Vector  
 US-09-935-194-1

Query Match 34.9%; Score 2672.4; DB 4; Length 7086;  
 Best Local Similarity 65.9%; Pred. No. 0; Mismatches 2071; Indels 312; Gaps 37;  
 Matches 4597; Conservative 0

QY 549 TTTTGAAGACCCACCTGTAGCTTTGGCAAGCTAGCTTAAGTAAGCCATTGCAAGG 608  
 Db 144 TTTTGAAGACCCACCCCGTAGG--TGGCAAGCTAGCTTAAGTAAGCCATTGCAAGG 201

QY 609 CATGG-AAAATACATTAATGAGATAGAGATCAAGTTAGAAACAGAGAGAC 667  
 Db 202 CATGGAAAAATACATTAATGAGATAGAGATCAAGTTAGAAACAGAGAGAC 261

QY 668 AGCAGATATGGCCCAACAGAGATATCTGTGTAGCAGTCTCTGCCCC--GCTCAGGGCC 726  
 Db 262 AGCTGAATA-CCAAACAGAGATATCTGTGTAGCAGTCTCTGCCCCCGCTCAGGGCC 317

QY 727 AAGAACAGTTGGAAACAGGAGAAAT-ATGGGCCAAACAGGATATCTGTGTAGCAGTTCT 785  
 Db 318 AAGAACAGATGAGACAGCTGAGTGATGGCCAAACAGGATATCTGTGTAGCAGTTCT 377

QY 786 GCGCCGGCTCAGGGCCAAACAGAGATGGTCCCGAGATGGTCCCGCCCTCAGCAGTTTC 845  
 Db 378 GCGCCGGCTCAGGGCCAAACAGAGATGGTCCCGAGATGGTCCCGCCCTCAGCAGTTTC 437

QY 846 TAGAGAACCATCAGATGTTTCAGGGTCCCAAGACCTG-ANATGACCTGTGCTTA 904  
 Db 438 TAGTGAATCATCAGATGTTTCAGGGTCCCAAGACCTG-ANATGACCTGTGCTTA 497

QY 905 TTTGAACTAACCAATCAGATTCGCTTCTCGCTTCTGTTCGCGGCTTCTGCTCCCGAGCT 964  
 Db 498 TTTGAACTAACCAATCAGATTCGCTTCTCGCTTCTGTTCGCGGCTTCTGCTCCCGAGCT 557

QY 965 CAATAAAGAGGCCAACACCCCTCACTCGGCGCGCAGTCTCCGATAGACTCGGTGCGC 1024  
 Db 558 CAATAAAGAGGCCAACACCCCTCACTCGGCGCGCAGTCTCCGATAGACTCGGTGCGC 617

QY 1025 CCGGTACCGGTATTCATTAAGCTCTTGTGTTGATCCGATCGTGAACCTCGCTG 1084  
 Db 618 CCGGTACCGGTATTCATTAAGCTCTTGTGTTGATCCGATCGTGAACCTCGCTG 677

QY 1085 ATCTTGGGAGGCTCTCTCAGATTGATGACTGCCACCTCGGGGCTCTTTCAATTGGA 1144

Db 678 TTCTTGGAGGGTCTCTCTGAGTGATTGACTACCCACGAGCGGGTCTTTCAATTGGG 737  
 QY 1145 GGTTCACCGAGATTTGGAGACCCCTGCCAGGACCAACCGACCCCGCGGGAGGTA 1204  
 Db 738 GGTCTGTCCGGATTTGGAGACCCCTGCCAGGACCAACCGA-CCACACCGGGAGGTA 796  
 QY 1205 AGCTGGCAGCAACTATCTGTGTCTGTCGAGTCTGTAGTCTGTATGACTGATTTATG 1264  
 Db 797 AGCTGGCAGCAACTATCTGTGTCTGTCGAGTCTGTAGTCTGTATGTTGATTTATG 856  
 QY 1265 GCGCTGCGTCTGACTAGTTAGTAACTAGTCTGTATCTGGCGACCCGTTGGTGAAC 916  
 Db 857 GCGCTGCGTCTGACTAGTTAGTAACTAGTCTGTATCTGGCGACCCGTTGGTGAAC 916  
 QY 1325 GACGAGTTTGGAGACCCCGCGGCAACCTGGGAGACGTCACGAGACTTCCGGGGCGG 1384  
 Db 917 GACGAGTTTGGAGACCCCGCGGCAACCTGGGAGACGTCACGAGACTTCCGGGGCGG 976  
 QY 1385 TTTTGTGCGCCGACCTGAGTCTCTAAATCCCGATCGTTTAGGACTCTTTGGTGACCC 1444  
 Db 977 TTTTGTGCGCCGACCTGAGGAGGAGTGGATCGATCGGATCCGACCCCGTC----- 1026  
 QY 1445 CTTAGAGGAGGATATGTGTTCTGTAGGAGACGAGAACTTAAACAGTTTCCCGCTCC 1504  
 Db 1027 -----AGGATATGTGTTCTGTAGGAGACGAGAACTTAAACAGTTTCCCGCTCC 1077  
 QY 1505 GTCTGAATTTTGTGTTTGGTGGACCGGACCGCGCGCGCTCTTGT---CTGCTG 1561  
 Db 1078 GTCTGAATTTTGTGTTTGGTGGACCGGACCGCGCGCTCTTGTCTGCTGACGCGTG 1137  
 QY 1562 CAGCATCTGTGTTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1621  
 Db 1138 CAGCATCTGTGTTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1197  
 QY 1622 CCGGGCTAGACTGTTTACCACCTCCCTTAAAGTTTGAACCTTAGTCTCACTGAAAGATG 1681  
 Db 1198 CC-----AGACTGTTTACCTCCCTTAAAGTTTGAACCTTAGTCTCACTGAAAGATG 1251  
 QY 1682 GCGGATCGCTCAACACAGTCTGATGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1741  
 Db 1252 GCGGATCGCTCAACACAGTCTGATGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1311  
 QY 1742 TGCAGATGGCCCAACTTTAACTCGGATGGCGGAGAGAGAGAGAGAGAGAGAGAGAG 1801  
 Db 1312 TGCAGATGGCCCAACTTTAACTCGGATGGCGGAGAGAGAGAGAGAGAGAGAGAGAG 1371  
 QY 1802 CATCACCCAGGTTAAGATCAAGGTTCTTTTACCTGGCCCGCATGACACCCAGAGAG 1861  
 Db 1372 CATCACCCAGGTTAAGATCAAGGTTCTTTTACCTGGCCCGCATGACACCCAGAGAG 1431  
 QY 1862 CCGCTTACATCTGAGCTGGAGAGCTTGGCTTTTGAACCCCTCTCTGGTCAAGCCCTT 1921  
 Db 1432 CCGCTTACATCTGAGCTGGAGAGCTTGGCTTTTGAACCCCTCTCTGGTCAAGCCCTT 1491  
 QY 1922 TGTACACCTTAAGCTCGGCT 1981  
 Db 1492 TGTACACCTTAAGCTCGGCT 1551  
 QY 1982 TCCTGTTTCAAGCCCGCTCTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2041  
 Db 1552 TCCTGTTTCAAGCCCGCTCTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1611  
 QY 2042 C-----CCATATGGCCATATGAGA 2061  
 Db 1612 CGGAATTCGATCTGATCAAGAGACAGGATGAGGATCGTTTGGCATGATTTGAACAGATG 1671  
 QY 2062 TCTTATATGGGACACCCCGCC-----CTTGTAACTTCCCTGACCTCTGACATGACAA 2115  
 Db 1672 GATTGACGAGGTTCTCGGCGGCTTGGTGGAGAGGCTATTCTGGCTATGATCGGAC 1731  
 QY 2116 GAGTTACTAAACAGCCCTCTCTCTCAAGCTCACTTACAGGCTTCTACTAGTCCAGACG- 2174

Db 1732 AACAGCAATCGCTGCTCTGATCGCGGTGTTCCGGCTGTACGCGAGGGGCCCGG 1791  
Qy 2175 -----AAGTCTGGAGACTCTGGCGGAGCCCTACAAAGAACTGGAGCCGCGTGG 2228  
Db 1792 TTCTTTTGTCAAGACCGACCTGTCCGGTCCCTGAATGAATGCTGAGGACGAGCGC 1851  
Qy 2229 TAC-----CTCACCTTACCGAGTCCGGGACACAGTGGGTC 2266  
Db 1852 GGTATCGTGGTGGCCAGCGGGCTTCTTGCAGCTGTCTCGACGTTGTCATG 1911  
Qy 2267 CGCCACACACAGACTAAGAACTAGAACTCGCTGGAAGAGACCTTACACAGTCTCTGTG 2326  
Db 1912 AAGCGGAGGAGCTGGTGTCTATTGGGGAGAGTCCCGGGCAGAGATCTCTGTATCTC 1971  
Qy 2327 ACCACCCCAACCGCCTCAAGTAGACGCGCATCGAGCTTGGATACACGCGCCCAAGTG 2386  
Db 1972 ACCTTGTCTC-CTGCCGAGAAAGTATCCATCATGTGGCTGATGCAATCGCGCGGCTGCATACG 2030  
Qy 2387 AAGGCTGCGACCCCGGGGTGGACCATCTCTAGACTGCCATGGATGGAGTGTATCA 2446  
Db 2031 CTTGAT-CCGGCTACTCCCATTCGACCAACCAAGCGAAATCGCATCGAGCGAGCAG 2089  
Qy 2447 TCCTCTTCTTG-----GTAGCAACAGCTACAGTGTCTCACTCCGACATCCAGCTGACCCA 2501  
Db 2090 TACTCGGATGGAGCGGTCTTGTCTGATCAGATGATCTGGACGAGAGCATCAGGGCT 2149  
Qy 2502 GAGCCCAAGACCTGAGCGCA-----GCGTGGGTGACAGAGTGACCATCACCT 2551  
Db 2150 CGCGCCAGCGAATGTTCCGAGGCTCAAGCGCGCATGCCGACGGGAGGATCTGT 2209  
Qy 2552 GTAAGGCAAGTCAGGATGGGTACTTCTGTAGCTTGGTACACAGCAGAACCCAGTAAG 2611  
Db 2210 CGTGACCCATGGCGATGCTGCTTCCGAATATCATGATGGTGGAAATGCGCGCTTCTCTG 2269  
Qy 2612 CTCRAAGCTGCTGATCTACTGGACAT--CCACCGGCGACACTGTGTGCCAAGCAGATT 2669  
Db 2270 ATTCACTACTGTGGCGCGGTGGGTGGCGGACGCTATCAGGACATAGCGTTGGCTAC 2329  
Qy 2670 CAGCGGTAGCGGTAGCGGTACCGACTTCACTTCACTACATCAGCAGCCTCCAGCCAGAGGA 2729  
Db 2330 CGGTATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCTGCTTTACGG 2389  
Qy 2730 CATGCCACCTACTACTGCCAATAATAGCCTCTATCGCTTATCGGTGTTGGCCCAAGGAGCAA 2789  
Db 2390 TATCGCGCTCCCGATTCCGAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTG 2449  
Qy 2790 GTGGAATC-----AAAGAGGTGGC 2811  
Db 2450 AGCGGACTCTGGGTTTCGAAATGACCGACCAAGCGACGCCCAACCTGCCATCAGAGAT 2509  
Qy 2812 TCAGGATCGGTGGATCCGGCTCTGGCTCAGGATCGAGGTCCAACCTGGTGAGAGC 2871  
Db 2510 TTCGATTTCACCGCGCCTTCTATGAAAGTTGGGCTTCGNAATCGTTTTCGGGAGCGCC 2569  
Qy 2872 GGTGGAGTGTGTCAACCTGGCGGTCCCTGCGCCTGTCTGCTCCGCACTTGGGTC 2931  
Db 2570 GGTGGATGATCTCCAGCGCGGGATCTCATGCTGGAGTTCTTCCGCCCAACCCCGGGCTC 2629  
Qy 2932 GATTTACCAATATGTGATGGTGGGTGAGACAGGACCTTGGAAAGTCTTGAGTGG 2991  
Db 2630 GATCCCTCGAGTTGGTTCA-----GCTGTGCTGAGGCTGGACGACCTCGCGGAGTT 2685  
Qy 2992 ATTGGAGAAATTCACAGATAGCAGTACGATTAACTATGCGCGCTCTCTAAAGATAGA 3051  
Db 2686 CTACCGGAGTCAAAATCTGTCGGCATCCAGGAAACCGAGCGGCTATCCGCGATCCA 2745  
Qy 3052 TTACATATATCGCGAGACAAACCCCAAGAAACATTTGTTCTGCAAAATGGACAGCCTGAGA 3111  
Db 2746 TGCCCCGAGACTCGAGAGTGGGAGGACGATGCGCGCTTTGGTCGAGGCGGATCTGCT 2805  
Qy 3112 CCCGAAGACACCGGGTCTATTTTGTGCAAGCCTTTACTCGGCTTCCCTGTTGCT 3171  
Db 2806 GGTCTTTCGCGCTCAGAAGCCATAGAGGCCACCGCATCCCGAGCATGCTGCTATTGTC 2865

Qy 3172 TATTGGGGCCAAAGGAGCCCGGTCACCGTCTCCAGTGTCTAAGCCACACAGACGCCAGG 3231  
Db 2866 TTCCCAATCTCCCGCTTGTCTGCTGCCACCCACCCACCCCAAGATAGAACTGACACT 2925  
Qy 3232 CCGGAGCA-----CCACACCGGGGCCACCATCGCTCGCAGCCCCCTGTCCTGC 3283  
Db 2926 ACTCAGCAATGCGATGCAATTTCTCATTTATTAGAAAGGACATGGAGTGGCACC 2985  
Qy 3284 GCCCAGAGCGGCTCGGCCAGCGGGGGCGCAGTGCACACGAGGGGGCTGGACTTCG 3343  
Db 2986 TTCCAGGCTCAAGGAAGCAGCGGGAGGGGCAACACAGATGGCTGCACTAGAGG 3045  
Qy 3344 CCCTGGATCCAAACTCTGTACTCTGTGTGATGAAATCCTTCTCATCTATGTGTGTCATC 3403  
Db 3046 CACAGTCAGGCTGATCAGCGAGCTCTAGCAATTTAGGTGACACTATAGATAGGGCCCTC 3105  
Qy 3404 TCACCTGCTTGTCTCTGAGTGAAGTTTCAGCAGGAGCGCAGAGCCCCCGGTACACAG 3463  
Db 3106 TAGATGATTAAGCTTAGATCTCGAGCGCGCCACCGGTCCGGATCCGAGCTCGGTAC-CA 3164  
Qy 3464 AGGCGCAGAAACAGCTCTATAAC---GAGCTCAATCTAGGACGAAGAGAGGATACGATG 3520  
Db 3165 AGCTTAAGTTCTCTCTATCACTGATAGGAAATCTCTATCACTGATAGGAGTCTTAT 3224  
Qy 3521 TTTTGGACAAGAGACGTGGCGGACCTGTGATGTTGGGGGAAAGCCGAGAAAGGAAGACC 3580  
Db 3225 ATATCTACTAGTCCGGATCACTATCACTGATAGGAAATCTCTATCACTGATAGGAGT 3284  
Qy 3581 CTCA--GGAAGGCTGTCAATGAACCTGCAGAAAGATAAGATGGGGAGGCTTACAGTGA 3638  
Db 3285 AGCATGATAGACTCCCAACCGTACACGCTACCGCCCATTTGCGTCAATGGGCGGAGT 3344  
Qy 3639 GATTGGGATGAAAGCGCAGCGCGGAGGGCAAGGGCAGCATGGCCTTTTACCAGGGTCT 3698  
Db 3345 TGTTACGACATTTTGGAAAGTCCCGTGTGATTTTGGTGCCTCAAAACAACTCCCATGAGCT 3404  
Qy 3699 CAGTACAGCCACCAAGGACACTTACGACGCCCTTTCACATGACGGCCCTGCCCCCTCGCTA 3758  
Db 3405 CAATGGGTGGAGACTTGGAAATCCCGTGAGTCAAAACCGCTATCCACGCCCATTTGATGT 3464  
Qy 3759 ACTCGACGCGCCCGGATCCGGATTAGTCAATTTGTTTAAAGACAGATATCAGTGTGTC 3818  
Db 3465 ACTGCCAAAACCGCATCACCATGGTAATAGCGATGACTTAATACGTAGATGACTGCCAAG 3524  
Qy 3819 CAGCTCTAGTTTGTACTCAACAATATACACAGCTGAAGCCTATAGAGTACGAGCCATAG 3878  
Db 3525 TAGG--AAAGTCCCATAGGTGATGTGCGGCAATATGCCAGGGGGCCATTTTACCCTC 3582  
Qy 3879 ATAAATAAAAGATTTTATTATTAGTCTCCAGAAAAGGGGGGAATGAAGACCCCACTGT 3938  
Db 3583 ATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGTATGTACTGCCAAGTGGGC 3642  
Qy 3939 AGGTTGGCAAGCTAGCTTAAGTAAGCCATTTTGAAGGCGATGGAAATATACATGTA 3998  
Db 3643 AGTTTACCGTAAATAGTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTTACTAT 3702  
Qy 3999 GAATAGAGAGTTCAGATCAAGTTTAGGAAACAGAGAGACACAGAAATATGGSCAAACAG 4058  
Db 3703 GGGAAACATACGTCAATTTTACGTCAATGGCGGGGGTTCGTTGGCGGTTCAGCCAGCGG 3762  
Qy 4059 GATATCTGTGTAGAGCAGTTCC-----TGCCCCGCTCAGGGGCCAAGAACAGTTTGAACA 4112  
Db 3763 GCCATTACCCTAGTTATGTAAACGGGAACCTCCATATATGGGTATGAACTAATGACCC 3822  
Qy 4113 GGAGAAATATGGCCAAACAGGATATCTGTGTAGAGAGTTCCTGCCCCGGCTCAGGGGCA 4172  
Db 3823 CGTAATTGATTACTATTAAATAACTAGTCAATAATCAATGTCAACGCGTATATCTGGCCG 3882  
Qy 4173 AGAAGATGTTCCCCAGATCGGTCCTCCGCTCAGCAGTTTCTAGA--GAACCATCAGAT 4231  
Db 3883 TACATCGGAAGCAGCGCAAAACCCCTAACCCCTAAGCAGATTTCTTATGCAATTTGTCGT 3942

QY 4232 GTTTCAGGGTCCCAAGAACCTGAAATGACCTGTGCTTATTTGAACTAAACCAATCA 4291  
 Db |||||  
 QY 3943 CAAAGCCTTGCTTGTGTAGCTTAAATTTTGTCTCGGCACCTACTCAGCGACCTCCAACAC 4002  
 Db |||||  
 QY 4292 GTTTCGCTTCTCGCTTCTGTTGCGGCGCTTCTGCTCCCGAGCTCAATAAAGAGCCCA 4351  
 Db |||||  
 QY 4003 ACAAGCAGGAGCAGATACCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGT 4062  
 Db |||||  
 QY 4352 ACCCTCACTCGGCGCGCAGTCTCTCCGATA-----GACTCGCTCGCCCGG 4398  
 Db |||||  
 QY 4063 CGACCATAGGGATCGGAGATCCGGGATATAATAAAGATTTTATTTAGTCTCCAGAA 4122  
 Db |||||  
 QY 4399 TACCCGTGTTCTCAATAAACCTCTTGCAGTTGCATCCGACTCGTGTCTCGCTGTTCT 4458  
 Db |||||  
 QY 4123 AAAGGGGGAATGAAGAAGCCACCTGTAGTTTGGCAAGCTAGCTTAAAGTAAGCCATT 4182  
 Db |||||  
 QY 4459 TGGAGGGTCTCTCTGAGTGAATGACTACCCGTACGCGGGTCT-----TTTCA 4506  
 Db |||||  
 QY 4183 TTGCAAGGCATGGAAAAATACATAA CTGAGAAATAGAGAAGTTTCAGATCAAGGTTCAGGAAC 4242  
 Db |||||  
 QY 4507 GTTTCCTCCACCTACACAGCTCTCACTAACTTCCTGATGTGCCGAGGACTCCGTCAG 4566  
 Db |||||  
 QY 4243 AGATGAAACAGCTGAATATGGGCAACAGGATATCTGTGGTAAGCAGTTCTTGCCCGG 4302  
 Db |||||  
 QY 4567 CCGG-----TTTTTGTATAATAAATGCAAGAACAGTGTTCCTTTCAAGCCA 4616  
 Db |||||  
 QY 4303 CTCAGGCCAAGACAGATGGAACAGCTGAATATGGCCAAACAGATATCTGTGGTAAG 4362  
 Db |||||  
 QY 4617 GACTACATCTGACTCTCGGCTTTATAAAGAAATGTTGAAGGCTCTGTGGAATATCTGTC 4676  
 Db |||||  
 QY 4363 CAGTTCTCGCCCGCTCAGGGCCAAAGAACAGATGCTCCAGATGCGGTCCAGCCCTCA 4422  
 Db |||||  
 QY 4677 CACACAGCTTTTAAAGATTTTATGCTCTCGATGAGGATTTAGTCAATATCTCTCG 4736  
 Db |||||  
 QY 4423 GCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCCAAGAGACCTGAATGACCTG 4482  
 Db |||||  
 QY 4737 TCTATTGCTGCTCTCTCGTATTATAAATTTCTAGTT-----TGACCTCCCTTCT 4789  
 Db |||||  
 QY 4483 TGCTTATTGAACTAAACCAATCAGTTCTGCTTCTGCTTCTGTTGCGGGCTTCTGCTCC 4542  
 Db |||||  
 QY 4790 GAGAGCAGGCGATTCAGAGTAGTTAAATCTGTGAGGCGAGCTTCTGTGAAAGGTTG 4849  
 Db |||||  
 QY 4543 CCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGCGCGCAGTCTCCGATTGACTG 4602  
 Db |||||  
 QY 4850 CTGGGCTCAGTGTGAGATTTTGGCATAA-----AAGGGTCTG 4890  
 Db |||||  
 QY 4603 AGTCCGCGGATCCCGTGTATCCAAATAAACCTCTTGAGTTGATCCGACTTGTGCTC 4662  
 Db |||||  
 QY 4891 CCGCTGTGTACAGACAGATCGGAATCTAGAGTGCACTCAGAGTCCCGCGGTTCT---- 4946  
 Db |||||  
 QY 4663 TGGCTGTTCTTGGAGGGTCTCTCTGAGTGATTTGACTTACCGGTACGCGGGGTTCTTC 4722  
 Db |||||  
 QY 4947 -----CGGGGCTCTGATCTCAGGGGATCTTTGCTPAGAGATCTCTACGCCGAGC 4997  
 Db |||||  
 QY 4723 ATTTGGGGGCTCGTCCGGATCGGAGACCCCTGCTCCAGGGACCAACCGACCCACACCG 4782  
 Db |||||  
 QY 4998 CATCTGGCGGGTACCGAGCTCGAATTCGTATCATGCT-CATAGCTGTTTCTGTGTG 5056  
 Db |||||  
 QY 4783 GAGGTAAAGTGTGCTCGCGGTTTCTGATGATGCGGTGAAACCTCTGACACATGCA 4842  
 Db |||||  
 QY 5057 AAATTGTATCTCGCTCACAATTCACAAATACAGAGCCGGAAGCATAAAGTGAAGC 5116  
 Db |||||  
 QY 4843 GCTCCGGAGACGGTACAGCTTGTCTGTAAAGGATGCGGGAGCAGCAAGCCGCTCA 4902  
 Db |||||  
 QY 5117 CTGGGTGCTTAATAGT-----GAGC 5138  
 Db |||||  
 QY 4903 GGGCGGCTCAGCGGGTGTGGGGGTGTCGGGGCGAGCCATGACCCAGTCAAGTAGCA 4962  
 Db |||||  
 QY 5139 TAACTACATTAATTCGCTTGGCTCAGTCCCGCTTCCAGTCGGGAAACCTGTCGTG 5198  
 Db |||||  
 QY 4963 TAGCGAGGTGATPACTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGGTGCA 5022  
 Db |||||  
 QY 5199 CAGTGCATTAAATGAATCGGCCAAACCGCGGGGAGAGGGGTTTGGGTATTGGGCGCTCT 5258  
 Db |||||

Db |||||  
 QY 5023 CATATGCGGTGTAATATACCGCACAGATCGGTAAAGAGAAAATACCGCATCAGGCGCTCT 5082  
 Db |||||  
 QY 5259 TCCGCTTCTCGCTCAGCTGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 5318  
 Db |||||  
 QY 5083 TCCGCTTCTCGCTCAGCTGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 5142  
 Db |||||  
 QY 5319 GCTCTCTCAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGACAGAAAGAAC 5378  
 Db |||||  
 QY 5143 GCTCACTCAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGACAGAAAGAAC 5202  
 Db |||||  
 QY 5379 ATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAAACCGTAAAGAGCGCGCTGTGCTGCGTTT 5438  
 Db |||||  
 QY 5203 ATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAAACCGTAAAGAGCGCGCTGTGCTGCGTTT 5262  
 Db |||||  
 QY 5439 TTCCATAGGCTCCGCGCCCTCGACAGCATCAAAAGATCGACGCTCAAGTCAGAGGTGG 5498  
 Db |||||  
 QY 5263 TTCCATAGGCTCCGCGCCCTCGACAGCATCAAAAGATCGACGCTCAAGTCAGAGGTGG 5322  
 Db |||||  
 QY 5499 CGAAACCCGACAGGACTATAAGATACAGAGGCTTTCCCTTGAAGCTCCCTCGTCGCG 5558  
 Db |||||  
 QY 5323 CGAAACCCGACAGGACTATAAGATACAGAGGCTTTCCCTTGAAGCTCCCTCGTCGCG 5382  
 Db |||||  
 QY 5559 TCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTGTCGCGCTTTCTCCCTTCGGGAAGC 5618  
 Db |||||  
 QY 5383 TCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTGTCGCGCTTTCTCCCTTCGGGAAGC 5442  
 Db |||||  
 QY 5619 GTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTGCGGTAGGTGCTGCTGCC 5678  
 Db |||||  
 QY 5443 GTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTGCGGTAGGTGCTGCTGCC 5502  
 Db |||||  
 QY 5679 AAGCTGGGCTGTGTCAGCAACCCCGCTTCCAGCCCGACCGCTGCGCTTATCCGCTAAC 5738  
 Db |||||  
 QY 5503 AAGCTGGGCTGTGTCAGCAACCCCGCTTCCAGCCCGACCGCTGCGCTTATCCGCTAAC 5562  
 Db |||||  
 QY 5739 TATCTGTTGAGTCCAAACCCCGTAAAGACACAGCTTATCCGCACTGCGCAGCAGCCACTG 5798  
 Db |||||  
 QY 5563 TATCTGTTGAGTCCAAACCCCGTAAAGACACAGCTTATCCGCACTGCGCAGCAGCCACTG 5622  
 Db |||||  
 QY 5799 AACAGATTAGCAGACGAGGTATGTAGCGGTGTCTACAGATTCTTGAAGTGTGGCT 5858  
 Db |||||  
 QY 5623 AACAGATTAGCAGACGAGGTATGTAGCGGTGTCTACAGATTCTTGAAGTGTGGCT 5682  
 Db |||||  
 QY 5859 AACTACGCTACACTAGAGGACAGTATTTGCTATCTGCTGCTGCTGCTGCTGCTGCTGCT 5918  
 Db |||||  
 QY 5683 AACTACGCTACACTAGAGGACAGTATTTGCTATCTGCTGCTGCTGCTGCTGCTGCTGCT 5742  
 Db |||||  
 QY 5919 TTCCGAAAAGAGTTGTTGAGTCTTTGATCCGCGCAACCAACCCGCTGCTGCTGCTGCTGCT 5978  
 Db |||||  
 QY 5743 TTCCGAAAAGAGTTGTTGAGTCTTTGATCCGCGCAACCAACCCGCTGCTGCTGCTGCTGCT 5802  
 Db |||||  
 QY 5979 TTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAGGATCTCAAGAGATCCTTTG 6038  
 Db |||||  
 QY 5803 TTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAGGATCTCAAGAGATCCTTTG 5862  
 Db |||||  
 QY 6039 ATCTTTTCTACGCGGTCTGACGCTCAGTGGAAACGAAACTCAGCTTAAGGGATTTTGGTC 6098  
 Db |||||  
 QY 5863 ATCTTTTCTACGCGGTCTGACGCTCAGTGGAAACGAAACTCAGCTTAAGGGATTTTGGTC 5922  
 Db |||||  
 QY 6099 ATGAGATTATCAAAAGGATCTTACCTAGATCTCTTTTAAATTTAAATTTAAATTTAAAT 6158  
 Db |||||  
 QY 5923 ATGAGATTATCAAAAGGATCTTACCTAGATCTCTTTTAAATTTAAATTTAAATTTAAAT 5982  
 Db |||||  
 QY 6159 TCAATCTAAAGTATATATGATGATACTTTGGTCTGACAGTTACCAATGCTTAATCAAGTGA 6218  
 Db |||||  
 QY 5983 TCAATCTAAAGTATATATGATGATACTTTGGTCTGACAGTTACCAATGCTTAATCAAGTGA 6042  
 Db |||||  
 QY 6219 GCACCTATCTCAGCGATCTGCTATTTGTTTCATCATAGTTGCTGCTGCTGCTGCTGCTGCT 6278  
 Db |||||  
 QY 6043 GCACCTATCTCAGCGATCTGCTATTTGTTTCATCATAGTTGCTGCTGCTGCTGCTGCTGCT 6102  
 Db |||||  
 QY 6279 TAGATAACTAGATACGAGGAGGCTTACCATCTGCGCCCGAGTGTGCTGAATGATACCGGA 6338  
 Db |||||

Db 6103 TAGATAACTACGATACGGAGGGCTTACCATCTGCCCCAGTGTGCAATGATACCGCGA 6162  
Qy 6339 GACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCGGAGCGGAG 6398  
Db 6163 GACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCGGAGCGGAG 6222  
Qy 6399 GCGAGAGTGTCTCGCAACTTATCCCGCTCCATCCAGTCTATTAATTTGTTGCGGAA 6458  
Db 6223 GCGAGAGTGTCTCGCAACTTATCCCGCTCCATCCAGTCTATTAATTTGTTGCGGAA 6282  
Qy 6459 GCTAGAGTAAGTAGTTCGCGAGTTAATAGTTTGGCCAAAGTTTGGCCATTTGCTACAGGC 6518  
Db 6283 GCTAGAGTAAGTAGTTCGCGAGTTAATAGTTTGGCCAAAGTTTGGCCATTTGCTACAGGC 6342  
Qy 6519 -TCGTGTTGTCACGCTCGTTCGTTGGTATGGCTTCATTCAGCTCCGGTTCCTCAACGATCA 6577  
Db 6343 ATCGTGGTGTACGCTCGTTCGTTGGTATGGCTTCATTCAGCTCCGGTTCCTCAACGATCA 6402  
Qy 6578 AGGCGAGTTACATGATCCCGCATGTTTGGCAAAAGCGGTTAGCTCTTCGTTCTCG 6637  
Db 6403 AGGCGAGTTACATGATCCCGCATGTTTGGCAAAAGCGGTTAGCTCTTCGTTCTCG 6462  
Qy 6638 ATCGTGTGTCAGAGTAAGTTGCGCGAGTTTATCACTCATGTTTATGCGAGCACTGCAT 6697  
Db 6463 ATCGTGTGTCAGAGTAAGTTGCGCGAGTTTATCACTCATGTTTATGCGAGCACTGCAT 6522  
Qy 6698 AATTCCTTACTGTCAATGCCATCCCGTAAGATGCTTTCTGTGACTGGTGTACTCAACC 6757  
Db 6523 AATTCCTTACTGTCAATGCCATCCCGTAAGATGCTTTCTGTGACTGGTGTACTCAACC 6582  
Qy 6758 AAGTCATTTGAGATAGTGTATGCGCGACCGAGTTGCTTTGCGCGCGGTCAATACGG 6817  
Db 6583 AAGTCATTTGAGATAGTGTATGCGCGACCGAGTTGCTTTGCGCGCGGTCAATACGG 6642  
Qy 6818 GATTAATCCGCGCCATAGCAGACTTTAAAGTGTCTCATCATTTGGAAACGTTCTCG 6877  
Db 6643 GATTAATCCGCGCCATAGCAGACTTTAAAGTGTCTCATCATTTGGAAACGTTCTCG 6702  
Qy 6878 GGGGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCATGTAAACCCACTCGT 6937  
Db 6703 GGGGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCATGTAAACCCACTCGT 6762  
Qy 6938 GCACCAACTGATCTTACGATCTTTTACTTTTACACGCGTTTCTGGGTGAGCAAAAACA 6997  
Db 6763 GCACCAACTGATCTTACGATCTTTTACTTTTACACGCGTTTCTGGGTGAGCAAAAACA 6822  
Qy 6998 GGAAGGCAAAATGCCGCAAAAAGGAATAGGGGACACGGAATGTTGAATCTCAT 7057  
Db 6823 GGAAGGCAAAATGCCGCAAAAAGGAATAGGGGACACGGAATGTTGAATCTCAT 6882  
Qy 7058 CTCTTCTTTTCAATATTTTGAAGCATTTTATCAGGGTTATGTTCTCATGAGCGGATAC 7117  
Db 6883 CTCTTCTTTTCAATATTTTGAAGCATTTTATCAGGGTTATGTTCTCATGAGCGGATAC 6942  
Qy 7118 ATATTGTAATGTTATGAAAATAAACAATAGGGGTTCCGCGCACATTTTCCCGGAAA 7177  
Db 6943 ATATTGTAATGTTATGAAAATAAACAATAGGGGTTCCGCGCACATTTTCCCGGAAA 7002  
Qy 7178 GTGCCACCTGAGCTTAAGAAACCAATTTATCATGACATTAACCTATAAAAATAGGCGT 7237  
Db 7003 GTGCCACCTGAGCTTAAGAAACCAATTTATCATGACATTAACCTATAAAAATAGGCGT 7062  
Qy 7238 ATCAGAGGCGCTTTCGTTCT 7257  
Db 7063 ATCAGAGGCGCTTTCGTTCT 7082

## RESULT 13

US-08-844-274-20/c

; Sequence 20, Application US/08844274B

; Patent No. 6218195

; GENERAL INFORMATION:

; APPLICANT: Fraser Jr., Malcom J.

; APPLICANT: Shirk, Paul D.  
; APPLICANT: Ellick, Teri A.  
; APPLICANT: Perera, Omaththage  
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System  
; TITLE OF INVENTION: for Insects  
; FILE REFERENCE: 0148.96  
; CURRENT APPLICATION NUMBER: US/08/844,274B  
; CURRENT FILING DATE: 1997-04-18  
; EARLIER APPLICATION NUMBER: 60/016,234  
; EARLIER FILING DATE: 1996-04-19  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.0 - beta  
; SEQ ID NO 20  
; LENGTH: 7560  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: p3E1.2hs/opd  
US-08-844-274-20

Query Match 34.4%; Score 2633.2; DB 3; Length 7560;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2646; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
Qy 5006 CCGGGTACCGAGCTCGAATTCGTAATCATGTGTCATAGCTGTTCTCTGTGTAATTTGTTA 5065  
Db 7413 CCGGGTACCGAGCTCGAATTCGTAATCATGTGTCATAGCTGTTCTCTGTGTAATTTGTTA 7354  
Qy 5066 TCCGCTCACAAATCCACACACATACGACGCGGAGCATAAAGTGAAGGCTCGGGTGC 5125  
Db 7353 TCCGCTCACAAATCCACACACATACGACGCGGAGCATAAAGTGAAGGCTCGGGTGC 7294  
Qy 5126 CTAATGAGTGAAGTAACTCACATTAATTCGCTTCGCTCACTGCGCGCTTTCCAGTCCGG 5185  
Db 7293 CTAATGAGTGAAGTAACTCACATTAATTCGCTTCGCTCACTGCGCGCTTTCCAGTCCGG 7234  
Qy 5186 AAACCTGTGTCAGCTGCAATTAATGAATCGGCAACGCGCGGAGAGCGGTTTGG 5245  
Db 7233 AAACCTGTGTCAGCTGCAATTAATGAATCGGCAACGCGCGGAGAGCGGTTTGG 7174  
Qy 5246 TATTGGGCGCTTTCGCGTCTTCGCTCACTGACTCGCTGCGCTGCGTTCGCTGCGTGC 5305  
Db 7173 TATTGGGCGCTTTCGCGTCTTCGCTCACTGACTCGCTGCGCTGCGTTCGCTGCGTGC 7114  
Qy 5306 GCGAGCGGTATCAGCTCACTCAAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGATAA 5365  
Db 7113 GCGAGCGGTATCAGCTCACTCAAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGATAA 7054  
Qy 5366 GCGAGGAAAGAACATGTGAGCAAAAGGCGCAAGAACCGCAAGAACCGTAAAAAGGCGCG 5425  
Db 7053 GCGAGGAAAGAACATGTGAGCAAAAGGCGCAAGAACCGCAAGAACCGTAAAAAGGCGCG 6994  
Qy 5426 GTTCTCGGCTTTTTCATAGGCTTCGCGCGCTTCGCGCGCTTCGCGCGCTTCGCGCTC 5485  
Db 6993 GTTCTCGGCTTTTTCATAGGCTTCGCGCGCTTCGCGCGCTTCGCGCGCTTCGCGCTC 6934  
Qy 5486 AAGTCAGAGTGGCGAAACCGCAGACGACTATAAAGATACCAAGCGTTTCCCTTGGAG 5545  
Db 6933 AAGTCAGAGTGGCGAAACCGCAGACGACTATAAAGATACCAAGCGTTTCCCTTGGAG 6874  
Qy 5546 CTCCTCTGCGCTTCTCTGTTCCGACCTTCGCGCTTACCGGATACCTGTCCGCTTTCT 5605  
Db 6873 CTCCTCTGCGCTTCTCTGTTCCGACCTTCGCGCTTACCGGATACCTGTCCGCTTTCT 6814  
Qy 5606 CCTTTCGGAAGCGTGGCGCTTTCTCATAGCTACGCTGTAGGTATCTCAGTTCTGGTGA 5665  
Db 6813 CCTTTCGGAAGCGTGGCGCTTTCTCATAGCTACGCTGTAGGTATCTCAGTTCTGGTGA 6754  
Qy 5666 GGTCTGCTCCCAAGCTGGGCTGTGTGCAAGACCCCGCTTCAGCCCGACCGCTGCGC 5725  
Db 6753 GGTCTGCTCCCAAGCTGGGCTGTGTGCAAGACCCCGCTTCAGCCCGACCGCTGCGC 6694  
Qy 5726 CTTATCCGGTAACCTATCGTCTTGAGTCCAAACCCCGGTAAAGACACGACTTTATCGCCACTGGC 5785

|||||  
6693 CTTATCCGGTAACATATCGTCTTGAGTCCAAACCGGTAAAGACAGCACTTATCGCCACTGGC 6634  
QY  
5786 AGCAGCCACTGGTAAACAGAGTATAGCAGCGAGGATATGTAGCGGTGCTACAGAGTTCTT 5845  
Db  
6633 AGCAGCCACTGGTAAACAGAGTATAGCAGCGAGGATATGTAGCGGTGCTACAGAGTTCTT 6574  
QY  
5846 GAAGTGGTGGCTTAACACAGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCT 5905  
Db  
6573 GAAGTGGTGGCTTAACACAGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCT 6514  
QY  
5906 GAAGCCAGTATACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACACCGC 5965  
Db  
6513 GAAGCCAGTATACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACACCGC 6454  
QY  
5966 TGGTAGCGGTGGTTTTTTTGGTTTGCAGCAGCAGATTAACCGCAGAAAAAAGGATCTCA 6025  
Db  
6453 TGGTAGCGGTGGTTTTTTTGGTTTGCAGCAGCAGATTAACCGCAGAAAAAAGGATCTCA 6394  
QY  
6026 AGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAAGCAAACTCAAGTTA 6085  
Db  
6393 AGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAAGCAAACTCAAGTTA 6334  
QY  
6086 AGGGAATTTTGGTCATGAGATTAACAAAAAGGATCTTCACTAGATCTCTTTAAATTTAAA 6145  
Db  
6333 AGGGAATTTTGGTCATGAGATTAACAAAAAGGATCTTCACTAGATCTCTTTAAATTTAAA 6274  
QY  
6146 ATGAAGTTTTAAATCAATCAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATG 6205  
Db  
6273 ATGAAGTTTTAAATCAATCAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATG 6214  
QY  
6206 CTTATACGTAGGACCTATCTCAGCGATCTGTCTATTTCTTATCTCATCATAGTTGCCGTG 6265  
Db  
6213 CTTATACGTAGGACCTATCTCAGCGATCTGTCTATTTCTTATCTCATCATAGTTGCCGTG 6154  
QY  
6266 ACTCCCGTCTGTAGATAAATACGATACGGAGGCTTACCATCTGGCCCCAGTGTGTC 6325  
Db  
6153 ACTCCCGTCTGTAGATAAATACGATACGGAGGCTTACCATCTGGCCCCAGTGTGTC 6094  
QY  
6326 AATGATACCGCAGACCCACGCTCACCGCTCACAGTTATCAGCAATAAACACGCCAGC 6385  
Db  
6093 AATGATACCGCAGACCCACGCTCACCGCTCACAGTTATCAGCAATAAACACGCCAGC 6034  
QY  
6386 CGGAAGGCGCAGCGCAGAGTGTCTCGCACTTTATCCGCTCCATCCAGTCTATTAA 6445  
Db  
6033 CGGAAGGCGCAGCGCAGAGTGTCTCGCACTTTATCCGCTCCATCCAGTCTATTAA 5974  
QY  
6446 TTGTTGCCGGGAAGCTAGAGTAAGTAGTTCCGCAAGTTAATAGTTTGGCAGACGTTGTC 6505  
Db  
5973 TTGTTGCCGGGAAGCTAGAGTAAGTAGTTCCGCAAGTTAATAGTTTGGCAGACGTTGTC 5914  
QY  
6506 CATTTGCTACAGGCTACGCTCGTGTGTTGGTATAGGCTTCAATCAGCTCCGG 6564  
Db  
5913 CATTTGCTACAGGCTACGCTCGTGTGTTGGTATAGGCTTCAATCAGCTCCGG 5854  
QY  
6565 TTCCCAACCATCAAGCGAGTTACATGATCCCCATGTTGTCGCAAAAAGCGGTAGCTC 6624  
Db  
5853 TTCCCAACCATCAAGCGAGTTACATGATCCCCATGTTGTCGCAAAAAGCGGTAGCTC 5794  
QY  
6625 CTTCCGCTCTCCGATCGTTTGTAGAGTAAGTTGGCCGAGTGTATCACTCATGTTAT 6684  
Db  
5793 CTTCCGCTCTCCGATCGTTTGTAGAGTAAGTTGGCCGAGTGTATCACTCATGTTAT 5734  
QY  
6685 GGAGCACTGCATTAATTTCTTACTGTCTATGCTCATGCCATCCGTAAGATGTTTTCTGTGACTGG 6744  
Db  
5733 GGAGCACTGCATTAATTTCTTACTGTCTATGCTCATGCCATCCGTAAGATGTTTTCTGTGACTGG 5674  
QY  
6745 TGAGTACTCAACCAAGTCAATCTGGAATAGTGTATGCGGACCGAGTGTCTTGCCC 6804  
Db  
5673 TGAGTACTCAACCAAGTCAATCTGGAATAGTGTATGCGGACCGAGTGTCTTGCCC 5614  
QY  
6805 GGCGTCAATACCGGATAATACCGCGCCACATAGCAAACTTTAAAGTCTCATCATTTGG 6864  
|||||

Db  
5613 GGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAAGTCTCATCATTTGG 5554  
QY  
6865 AAAAGTTTCTTCGGGGCGAAAACCTCTCAAGGATCTTACGGCTGTGTAGATCCAGTTGGAT 6924  
Db  
5553 AAAAGTTTCTTCGGGGCGAAAACCTCTCAAGGATCTTACGGCTGTGTAGATCCAGTTGGAT 5494  
QY  
6925 GTAACCCCACTCGTGACCACTGATCTTACGATCTTCTTACTTTTCAACAGCGTTTCTGG 6984  
Db  
5493 GTAACCCCACTCGTGACCACTGATCTTACGATCTTCTTACTTTTCAACAGCGTTTCTGG 5434  
QY  
6985 GTGACAAAAACAGGAAGCAAAATGCCGCAAAAAAGGAAATAAGGGCGACACGAAATG 7044  
Db  
5433 GTGACAAAAACAGGAAGCAAAATGCCGCAAAAAAGGAAATAAGGGCGACACGAAATG 5374  
QY  
7045 TTGAATATCTCATCTCTTCTTCTTCTTCAATATTTTGAAGCATTTTATCAGGGTATTTGCT 7104  
Db  
5373 TTGAATATCTCATCTCTTCTTCTTCTTCAATATTTTGAAGCATTTTATCAGGGTATTTGCT 5314  
QY  
7105 CATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAAACAAATAGGGGTTCCCGGCAC 7164  
Db  
5313 CATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAAACAAATAGGGGTTCCCGGCAC 5254  
QY  
7165 ATTTCCCGAAAAAGTGCACCTGACGTCTAAGAAAAACATTTATCATGACATTAACCTTA 7224  
Db  
5253 ATTTCCCGAAAAAGTGCACCTGACGTCTAAGAAAAACATTTATCATGACATTAACCTTA 5194  
QY  
7225 TAAAAATAGGCGTATCAGAGGCCCTTTCGTCTCGCGGCTTTCGTGTGATGACGGTGAATA 7284  
Db  
5193 TAAAAATAGGCGTATCAGAGGCCCTTTCGTCTCGCGGCTTTCGTGTGATGACGGTGAATA 5134  
QY  
7285 CCTCTGACACATGACAGCTCCCGGAGACGCTCACAGCTTGTCTGTAAAGCGGATGCCGGAG 7344  
Db  
5133 CCTCTGACACATGACAGCTCCCGGAGACGCTCACAGCTTGTCTGTAAAGCGGATGCCGGAG 5074  
QY  
7345 CAGACAACCCGCTCAGGGCGGCTCAGCGGCTGTGCGGGGTGTCGGGGCTGCGCTTAACTTA 7404  
Db  
5073 CAGACAACCCGCTCAGGGCGGCTCAGCGGCTGTGCGGGGTGTCGGGGCTGCGCTTAACTTA 5014  
QY  
7405 TCGGCATCAGAGCAGATTTACTGAGAGTGCACCATATGCGGTGTGAATACCCGACAG 7464  
Db  
5013 TCGGCATCAGAGCAGATTTACTGAGAGTGCACCATATGCGGTGTGAATACCCGACAG 4954  
QY  
7465 ATGCGTAAGGAGAAAAATACCGCATCAGCGCCATTCGCGCATTTACGCTGCGCAACTGTGTG 7524  
Db  
4953 ATGCGTAAGGAGAAAAATACCGCATCAGCGCCATTCGCGCATTTACGCTGCGCAACTGTGTG 4894  
QY  
7525 GGAAGGGCGATCGGTGCGGGCTCTTCGCTATTTACGCCAGCTGGCGAAAGGGGGATGTGC 7584  
Db  
4893 GGAAGGGCGATCGGTGCGGGCTCTTCGCTATTTACGCCAGCTGGCGAAAGGGGGATGTGC 4834  
QY  
7585 TGCAAGGCGATTAAGTTGGGTAAACGCCAGGGTTTTCCAGTCACGACCTTGTAAACGAC 7644  
Db  
4833 TGCAAGGCGATTAAGTTGGGTAAACGCCAGGGTTTTCCAGTCACGACCTTGTAAACGAC 4774  
QY  
7645 GSCCAGTGCC 7654  
Db  
4773 GSCCAGTGCC 4764

## RESULT 14

US-09-598-421-20/c

; Sequence 20, Application US/09598421

; Patent No. 6551825

; GENERAL INFORMATION:

; APPLICANT: Fraser Jr., Malcom J.

; APPLICANT: Shirk, Paul D.

; APPLICANT: Elick, Teri A.

; APPLICANT: Perera, Onaththage

; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System

; FILE OF INVENTION: for Insects

; FILE REFERENCE: 0148.96

; CURRENT APPLICATION NUMBER: US/09/598,421

; CURRENT FILING DATE: 2000-06-19



```
; PRIOR APPLICATION NUMBER: 60/016,234
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 20
; LENGTH: 7560
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: p3E1.2hs/opd
US-09-598-421-20

Query Match      34.4%; Score 2633.2; DB 4; Length 7560;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2646; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 5006 CCGGTACCGAGCTCGAATTCGTAATCATGGTCATAGCTGTTCCCTGCTGTGTAATTTGTTA 5065
DB 7413 CCGGTACCGAGCTCGAATTCGTAATCATGGTCATAGCTGTTCCCTGCTGTGTAATTTGTTA 7354
QY 5066 TCCGCTCAAAATCCACACAAATACGAGCCGGAAGCATAAAGTGTAAAGCCCTGGGGTGC 5125
DB 7353 TCCGCTCAAAATCCACACAAATACGAGCCGGAAGCATAAAGTGTAAAGCCCTGGGGTGC 7294
QY 5126 CTAATGAGTGAGCTAACTCACAATTAATTCGGTTGGCTCCTACGCGCGCTTTCCAGTCGGG 5185
DB 7293 CTAATGAGTGAGCTAACTCACAATTAATTCGGTTGGCTCCTACGCGCGCTTTCCAGTCGGG 7234
QY 5186 AAACCTGTCGCGAGCTGCAATTAATGAATCGGCCAAGCGCGGGAGAGCGGTTTCGG 5245
DB 7233 AAACCTGTCGCGAGCTGCAATTAATGAATCGGCCAAGCGCGGGAGAGCGGTTTCGG 7174
QY 5246 TATTGGGCGCTCTTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGCTCGCTCGCTCGG 5305
DB 7173 TATTGGGCGCTCTTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGCTCGCTCGG 7114
QY 5306 GCGAGCGGTATCAGTCTCACTAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAA 5365
DB 7113 GCGAGCGGTATCAGTCTCACTAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAA 7054
QY 5366 GCGAGGAAGACATGTGAGCAAAAGCCGCAAAAGCCGGAAGCCGTAAGAGGCGGC 5425
DB 7053 GCGAGGAAGACATGTGAGCAAAAGCCGCAAAAGCCGGAAGCCGTAAGAGGCGGC 6994
QY 5426 GTTGCTGCGGTTTTTCCATAGGCTCCGCGCCCTCGTACGAGCATCACAAAATCAGACGCTC 5485
DB 6993 GTTGCTGCGGTTTTTCCATAGGCTCCGCGCCCTCGTACGAGCATCACAAAATCAGACGCTC 6934
QY 5486 AAGTCAGAGTGGCGGAAACCCGACAGGACTATAAGATACAGAGGCTTTCCCTCGGAAG 5545
DB 6933 AAGTCAGAGTGGCGGAAACCCGACAGGACTATAAGATACAGAGGCTTTCCCTCGGAAG 6874
QY 5546 CTCCTCTGCGCTCTCTGTTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTCT 5605
DB 6873 CTCCTCTGCGCTCTCTGTTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTCT 6814
QY 5606 CCCTTCGCGAAGCGTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTA 5665
DB 6813 CCCTTCGCGAAGCGTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTA 6754
QY 5666 GGTGCTTCGCTCCAGTCTGGGCTGTGTGACAGAACCCCGCTTACGCGCCGACCGCTGCGC 5725
DB 6753 GGTGCTTCGCTCCAGTCTGGGCTGTGTGACAGAACCCCGCTTACGCGCCGACCGCTGCGC 6694
QY 5726 CTTATCCCGTAACTATCTGTTAGTCCAAACCGGTAAAGACAGCACTTATCCGCACTGGC 5785
DB 6693 CTTATCCCGTAACTATCTGTTAGTCCAAACCGGTAAAGACAGCACTTATCCGCACTGGC 6634
QY 5786 AGCAGCCACTGTAAACAGATTTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTT 5845
DB 6633 AGCAGCCACTGTAAACAGATTTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTT 6574
QY 5846 GAAAGTGGTGGCTTAACGCGCTACACTAGAGGACAGTATTTGGTATCTCGGCTCTGCT 5905
```

```
DB 6573 GAAGTGGTGGCTTAACCTACGGCTACACTAGAAGGACAGTATTTGGTATCTCGGCTCTGCT 6514
QY 5906 GAAGCCAGTCTACCTTCGGAAAAAGAGTTGGTAGCTCTTTGATCCGGCAAAACCAACCCAGC 5965
DB 6513 GAAGCCAGTCTACCTTCGGAAAAAGAGTTGGTAGCTCTTTGATCCGGCAAAACCAACCCAGC 6454
QY 5966 TGGTAGCGGTGGTTTTTTTGTTCGAAGCAGCAGANTTACGCGCAAGAAAAAGAGTCTCA 6025
DB 6453 TGGTAGCGGTGGTTTTTTTGTTCGAAGCAGCAGANTTACGCGCAAGAAAAAGAGTCTCA 6394
QY 6026 AGAAGATCCTTTGATCTTTTCTACGGGCTCTGACGCTCAGTGAACGAAAACTCAGTTA 6085
DB 6393 AGAAGATCCTTTGATCTTTTCTACGGGCTCTGACGCTCAGTGAACGAAAACTCAGTTA 6334
QY 6086 AGGATTTTGGTCTGATGAGATTATCAAAAAGGATCTTACCTAGATCTCTTTTAAATTTAAA 6145
DB 6333 AGGATTTTGGTCTGATGAGATTATCAAAAAGGATCTTACCTAGATCTCTTTTAAATTTAAA 6274
QY 6146 ATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTTGGTCTGACAGTTACCAATG 6205
DB 6273 ATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTTGGTCTGACAGTTACCAATG 6214
QY 6206 CTTAATCAGTCAGGACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCATAGTTGCTG 6265
DB 6213 CTTAATCAGTCAGGACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCATAGTTGCTG 6154
QY 6266 ACTCCCGTCTGTAGATAACTACGATACGGAGGGCTTACCATCTGCGCCAGTGCTGC 6325
DB 6153 ACTCCCGTCTGTAGATAACTACGATACGGAGGGCTTACCATCTGCGCCAGTGCTGC 6094
QY 6326 AATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTTATCAGCAATAAACCCAGCCAGC 6385
DB 6093 AATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTTATCAGCAATAAACCCAGCCAGC 6034
QY 6386 CGGAAGGGCCGAGCGCAGAGTGTCTGCAACTTTTATCCGCTCCATCCAGTCTATTAA 6445
DB 6033 CGGAAGGGCCGAGCGCAGAGTGTCTGCAACTTTTATCCGCTCCATCCAGTCTATTAA 5974
QY 6446 TTGTTGCGGGAAGCTAGAGTAAGTTCGCCAGTTAAATAGTTTGGCCCAAGTTGTTGC 6505
DB 5973 TTGTTGCGGGAAGCTAGAGTAAGTTCGCCAGTTAAATAGTTTGGCCCAAGTTGTTGC 5914
QY 6506 CATTGCTACAGGC-TCGTGCTGTACGCTCGCTGTTGGTATGGCTTCATTTCAGTCCGG 6564
DB 5913 CATTGCTACAGGCATCGTGGTGTTCAGCTCGTGGTATGGTATGGCTTCATTTCAGCGCGCG 5854
QY 6565 TTCCCAACGATCAAGCGGAGTTAATGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTC 6624
DB 5853 TTCCCAACGATCAAGCGGAGTTAATGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTC 5794
QY 6625 CTTCCGCTCCCGATCGTTGTGCAAGTAAGTTCGCCAGTGTATCAGTCACTCATGTTAT 6684
DB 5793 CTTCCGCTCCCGATCGTTGTGCAAGTAAGTTCGCCAGTGTATCAGTCACTCATGTTAT 5734
QY 6685 GGCAAGCATGCTAATTTCTTTACTGTCTATGCCATCCGTAAGATGCTTTCTGTGACTGG 6744
DB 5733 GGCAAGCATGCTAATTTCTTTACTGTCTATGCCATCCGTAAGATGCTTTCTGTGACTGG 5674
QY 6745 TGAGTATCAACCAAGTCAATTCAGGAATAGTGTATGCGGCGACCGAGTGTCTCTTGGCC 6804
DB 5673 TGAGTATCAACCAAGTCAATTCAGGAATAGTGTATGCGGCGACCGAGTGTCTCTTGGCC 5614
QY 6805 GGCGTCAATACGGGATTAATACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCTGG 6864
DB 5613 GGCGTCAATACGGGATTAATACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCTGG 5554
QY 6865 AAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTGTAGATCCAGTTCGAT 6924
DB 5553 AAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTGTAGATCCAGTTCGAT 5494
QY 6925 GTAAACCACTCGTGACCCCACTGATCTTACAGCATCTTTTACTTTTCAACGAGTTTCTGG 6984
```



Db 5493 GTAACCCACTGCTGACCAACAGTATCTTCACGATCTTTACTTTTACAGCGGTTTCTGG 5434  
Qy 6985 GTGAGCAAAAACAGGAAGCAAAAATGCCGCAAAAAGGAATAAGGGCGACACGAAATG 7044  
Db 5433 GTGAGCAAAAACAGGAAGCAAAAATGCCGCAAAAAGGAATAAGGGCGACACGAAATG 5374  
Qy 7045 TTGAATACATACACTCTTCTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCT 7104  
Db 5373 TTGAATACATACACTCTTCTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCT 5314  
Qy 7105 CATGAGCGATACATATTGAATGATTTAGAAAATAAACAATAAGGGTTCCGCGAC 7164  
Db 5313 CATGAGCGATACATATTGAATGATTTAGAAAATAAACAATAAGGGTTCCGCGAC 5254  
Qy 7165 ATTTCGCCAAAAGTCCACCTGACGCTCTAAGAAACCAATTATTATCATGACATTAACCTA 7224  
Db 5253 ATTTCGCCAAAAGTCCACCTGACGCTCTAAGAAACCAATTATTATCATGACATTAACCTA 5194  
Qy 7225 TAAATAAGCGGTATCAGAGGCCCTTTCGTCTCGCGGCTTTTCGGTGATGACGGTGAATA 7284  
Db 5193 TAAATAAGCGGTATCAGAGGCCCTTTCGTCTCGCGGCTTTTCGGTGATGACGGTGAATA 5134  
Qy 7285 CCTCTGACATGACGCTCCCGAGACGCTCAGAGCTTGTCTGTAAGCGGATGCCGGAG 7344  
Db 5133 CCTCTGACATGACGCTCCCGAGACGCTCAGAGCTTGTCTGTAAGCGGATGCCGGAG 5074  
Qy 7345 CAGACAAGCCGTCAGGCGGCTGTCAGCGGCTGTGCGGGTGTGCGGGTGCCTTAACCTA 7404  
Db 5073 CAGACNAGCCGTCAGGCGGCTGTCAGCGGCTGTGCGGGTGTGCGGGTGCCTTAACCTA 5014  
Qy 7405 TCGGCATCAGAGCAGATTTGTAAGAGTGCACCATATGCGGTGTGAAATACCGCACAG 7464  
Db 5013 TCGGCATCAGAGCAGATTTGTAAGAGTGCACCATATGCGGTGTGAAATACCGCACAG 4954  
Qy 7465 ATGCGTAAGAGAAAATACCGATCAGGCGCATTCGCCCATTCAGGCTCGGCAACTGTG 7524  
Db 4953 ATGCGTAAGAGAAAATACCGATCAGGCGCATTCGCCCATTCAGGCTCGGCAACTGTG 4894  
Qy 7525 GGAAGGGCATCGTTCGCGGCTCTTCGCTATTAGCCAGCTGGCGAAGGGGATGTGC 7584  
Db 4893 GGAAGGGCATCGTTCGCGGCTCTTCGCTATTAGCCAGCTGGCGAAGGGGATGTGC 4834  
Qy 7585 TGCAAGGCGATTAAGTTGGTTAAGCGCGAGGTTTCCAGTCAAGCGCTGTAAGAACGAC 7644  
Db 4833 TGCAAGGCGATTAAGTTGGTTAAGCGCGAGGTTTCCAGTCAAGCGCTGTAAGAACGAC 4774  
Qy 7645 GGCCAGTGCC 7654  
Db 4773 GGCCAGTGCC 4764

RESULT 15  
US-09-205-817A-2  
; Sequence 2, Application US/09205817A  
; Patent No. 6642028  
; GENERAL INFORMATION:  
; APPLICANT: Ill, Charles R. et al.  
; TITLE OF INVENTION: NOVEL VECTORS AND GENES EXHIBITING  
; TITLE OF INVENTION: INCREASED EXPRESSION  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD, LLP  
; STREET: 28 STATE STREET  
; CITY: BOSTON  
; STATE: MASSACHUSETTS  
; COUNTRY: US  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/205,817A  
; FILING DATE: 04 DECEMBER 1998  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/067,614  
; FILING DATE: 05 DECEMBER 1997  
; APPLICATION NUMBER: US 60/071,596  
; FILING DATE: 16 JANUARY 1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: REMILLARD, JANE E.  
; REGISTRATION NUMBER: 38,872  
; REFERENCE/DOCKET NUMBER: TTI-180  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)742-4214  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9164 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1006..5376  
; US-09-205-817A-2  
  
Query Match 34.3%; Score 2623.4; DB 4; Length 9164;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2635; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
  
Qy 5019 TCGAATTCGTATCATGTGTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCAAAAT 5078  
Db 6513 TTGAATTCGTATCATGTGTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCAAAAT 6572  
Qy 5079 CCACACACATACAGAGCCGGAAGCATAAAGTGTAAAGCTGGGGTGCCTATGATGAGC 5138  
Db 6573 CCACACACATACAGAGCCGGAAGCATAAAGTGTAAAGCTGGGGTGCCTATGATGAGC 6632  
Qy 5139 TAATCATCATTAATTCGCTGCGCTCACTGCCGCTTCCAGTCCGGGAAACCTGCTGTC 5198  
Db 6633 TAATCATCATTAATTCGCTGCGCTCACTGCCGCTTCCAGTCCGGGAAACCTGCTGTC 6692  
Qy 5199 CAGCTGCAATTAATGAATCGGCAACCGCGGGGAGAGCGGTTTCGCTATTGGGCGCTCT 5258  
Db 6693 CAGCTGCAATTAATGAATCGGCAACCGCGGGGAGAGCGGTTTCGCTATTGGGCGCTCT 6752  
Qy 5259 TCCGCTTCTCGCTCACTGACTCGCTGCGCTGCGCTGCGCTGCGGCGAGCGGTATCA 5318  
Db 6753 TCCGCTTCTCGCTCACTGACTCGCTGCGCTGCGCTGCGCTGCGGCGAGCGGTATCA 6812  
Qy 5319 GCTCACTCAAAGCGGTAATACGGTTATCCACAGATCAGGGGATAACGAGGAAGAC 5378  
Db 6813 GCTCACTCAAAGCGGTAATACGGTTATCCACAGATCAGGGGATAACGAGGAAGAC 6872  
Qy 5379 ATGTGAGCAAAAGGCGCAGCAAAAGCCAGGAACCCGTAAAGAGCCGCGTGTGCGGTTT 5438  
Db 6873 ATGTGAGCAAAAGGCGCAGCAAAAGCCAGGAACCCGTAAAGAGCCGCGTGTGCGGTTT 6932  
Qy 5439 TTCCATAGGCTCCGCCCGCTGACGAGCATCAAAAATTCAGCGCTCAAGTCAGAGGTGG 5498  
Db 6933 TTCCATAGGCTCCGCCCGCTGACGAGCATCAAAAATTCAGCGCTCAAGTCAGAGGTGG 6992  
Qy 5499 CGAAACCCGACAGGACTATAAGATACAGGGGTTTCCCGCTGGAGCTCCCTCGCTGCGC 5558  
Db 6993 CGAAACCCGACAGGACTATAAGATACAGGGGTTTCCCGCTGGAGCTCCCTCGCTGCGC 7052  
Qy 5559 TCTCTGTTCGACCCCTGCGCTTACCGATACCTGTCGCGCTTCTCCCTTCGGGAAGC 5618  
Db 7053 TCTCTGTTCGACCCCTGCGCTTACCGATACCTGTCGCGCTTCTCCCTTCGGGAAGC 7112  
Qy 5619 GTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGTTAGGTTCGCTCC 5678

Db	7113	GTGGCGCTTCTCTAAGCTCAGCTCAGCTGTAGGTATCTCAGTTCCGTTAGGTTCGTTTCGCTCC	7117
Qy	5679	AAGCTGGCGCTGTGTGCACGAACCCCGCTTCAGCGCCAGCCGCTTCGCGCTTTATCCGGTAAC	5738
Db	7173	AAGCTGGCGCTGTGTGCACGAACCCCGCTTCAGCGCCAGCCGCTTCGCGCTTTATCCGGTAAC	7232
Qy	5739	TATCGTCTTGAGTCCAAACCGGTAAGAACAAGACTTATTCGCCACTGGCAGCAGCCACTGGT	5798
Db	7233	TATCGTCTTGAGTCCAAACCGGTAAGAACAAGACTTATTCGCCACTGGCAGCAGCCACTGGT	7292
Qy	5799	AACAGGATTAAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTTGAAGTGGTGGCCT	5858
Db	7293	AACAGGATTAAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTTGAAGTGGTGGCCT	7352
Qy	5859	AACTACCGCTACACTAGAAGACAGTATTTGGTATCTCGCGCTCTGCTGAAGCCAGTTTACC	5918
Db	7353	AACTACCGCTACACTAGAAGACAGTATTTGGTATCTCGCGCTCTGCTGAAGCCAGTTTACC	7412
Qy	5919	TTCCGAAAAAGATTGGTAGCTCTTGATCCGGCAAAACAAACCAACCGCTGTGTAGCGGTGGT	5978
Db	7413	TTCCGAAAAAGATTGGTAGCTCTTGATCCGGCAAAACAAACCAACCGCTGTGTAGCGGTGGT	7472
Qy	5979	TTTTTTCTTGGCAGCAGCAGATTAACCGCAGAAAAAAGAGTCTCAAGAGATCTCTTTG	6038
Db	7473	TTTTTTCTTGGCAGCAGCAGATTAACCGCAGAAAAAAGAGTCTCAAGAGATCTCTTTG	7532
Qy	6039	ATCTTTTCTACGGGCTCTGACGCTCAGTGGAAACGAAAACTCACGTTAAGGGATTTTGGTC	6098
Db	7533	ATCTTTTCTACGGGCTCTGACGCTCAGTGGAAACGAAAACTCACGTTAAGGGATTTTGGTC	7592
Qy	6099	ATGAGATTATCAAAAAGGATCTTCACCTAGATCCCTTTTAAATTAATAATGAAGTTTAAA	6158
Db	7593	ATGAGATTATCAAAAAGGATCTTCACCTAGATCCCTTTTAAATTAATAATGAAGTTTAAA	7652
Qy	6159	TCATCTAAAGTATATATGAGTAAACTTGGTCTGACAGATTACCAATGCTTTAATCAGTGAG	6218
Db	7653	TCATCTAAAGTATATATGAGTAAACTTGGTCTGACAGATTACCAATGCTTTAATCAGTGAG	7712
Qy	6219	GCACCTATCTCAGCGATCTGTCTATTTTGGTTTCACTCCATAGTTGCTGACCTCCCGCTCGTG	6278
Db	7713	GCACCTATCTCAGCGATCTGTCTATTTTGGTTTCACTCCATAGTTGCTGACCTCCCGCTCGTG	7772
Qy	6279	TAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGCGA	6338
Db	7773	TAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGCGA	7832
Qy	6339	GACCCAGCTCACCGGCTCCAGATTATCAGCAATAAACACAGCAGCCGGAAGGGCCGAG	6398
Db	7833	GACCCAGCTCACCGGCTCCAGATTATCAGCAATAAACACAGCAGCCGGAAGGGCCGAG	7892
Qy	6399	CGCAGAAGTGGTCTCTGCAACTTTTATCCGCTCCATCCAGTCTAATTAATTTGTCGCCGGAA	6458
Db	7893	CGCAGAAGTGGTCTCTGCAACTTTTATCCGCTCCATCCAGTCTAATTAATTTGTCGCCGGAA	7952
Qy	6459	GCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGGCAACGTTGTGTCATTGCTCAGAGC	6518
Db	7953	GCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGGCAACGTTGTGTCATTGCTCAGAGC	8012
Qy	6519	-TGTTGTTGTCACGCTCGTGTGGTGGTATCGGCTTCATTCAAGTCCGGTTCCCAACATCA	6577
Db	8013	ATCGTGGTGTACGCTCGTGTGGTGGTATCGGCTTCATTCAAGTCCGGTTCCCAACATCA	8072
Qy	6578	AGCGAGTTACATGATCCCCATGTTGTGCAAAAAGCGGTAGCTCTTCGCTCCCTCCG	6637
Db	8073	AGCGAGTTACATGATCCCCATGTTGTGCAAAAAGCGGTAGCTCTTCGCTCCCTCCG	8132
Qy	6638	ATCGTGTTCAGAGTAAGTTTGGCGCAGTGTATCACTCATGGTTATGGCAGCACTGCAT	6697
Db	8133	ATCGTGTTCAGAGTAAGTTTGGCGCAGTGTATCACTCATGGTTATGGCAGCACTGCAT	8192
Qy	6698	AATTCCTTACTGTATGCCATCCGTAAAGATCTTTTCTGTGACTGGTGAGTACTCAACC	6757
Db	8193	AATTCCTTACTGTATGCCATCCGTAAAGATCTTTTCTGTGACTGGTGAGTACTCAACC	8252

QY	6758	AAGTCATTCTGAGAAATAGTGTATGTCGGCGGACCGAGTGTCTCTTGGCCGGGTCAATACGG	6817
Db	8253	AAGTCATTCTGAGAAATAGTGTATGTCGGCGGACCGAGTGTCTCTTGGCCGGGTCAATACGG	8312
QY	6818	GATAAATACCGGCCACATACAGAACTTTAAAGTGTCTCATCATCGGAACCTTCTTCG	6877
Db	8313	GATAAATACCGGCCACATACAGAACTTTAAAGTGTCTCATCATCGGAACCTTCTTCG	8372
QY	6878	GGCGGAAAACTCTCAAGGATCTTACCGCTGTTCAGATCCAGTTTCGATGTAAACCCACTCGT	6937
Db	8373	GGCGGAAAACTCTCAAGGATCTTACCGCTGTTCAGATCCAGTTTCGATGTAAACCCACTCGT	8432
QY	6938	GCACCCAACTGATCTTCAGCATCTTTTACATTTTCAACAGCGTTTCTGGGTGAGCAAAACA	6997
Db	8433	GCACCCAACTGATCTTCAGCATCTTTTACATTTTCAACAGCGTTTCTGGGTGAGCAAAACA	8492
QY	6998	GGAAGGCAAAATGCCGCAAAAAAGGAAATGAAGGCGACACGGAATGTTCGATCTCATTA	7057
Db	8493	GGAAGGCAAAATGCCGCAAAAAAGGAAATGAAGGCGACACGGAATGTTCGATCTCATTA	8552
QY	7058	CTCTTCCTTTTTCAATAATTATGAAGCATTTATCAGGTTTATTGTCTCATGACGGGATC	7117
Db	8553	CTCTTCCTTTTTCAATAATTATGAAGCATTTATCAGGTTTATTGTCTCATGACGGGATC	8612
QY	7118	ATAATTTGAATGTATTAGAAAAATAAACAAATAGGGGTTCCCGCGCACATTTCCCGGAAAA	7177
Db	8613	ATAATTTGAATGTATTAGAAAAATAAACAAATAGGGGTTCCCGCGCACATTTCCCGGAAAA	8672
QY	7178	GTGCCACTGACGCTTAAGAAAACTATTATCATGACATTAAACCTTATAAAATAGGCGT	7237
Db	8673	GTGCCACTGACGCTTAAGAAAACTATTATCATGACATTAAACCTTATAAAATAGGCGT	8732
QY	7238	ATCACGAGGCCCTTCGTCTCGCGCGTTTCGCGTATGACGGTCAAAACCTCTGACACATG	7297
Db	8733	ATCACGAGGCCCTTCGTCTCGCGCGTTTCGCGTATGACGGTCAAAACCTCTGACACATG	8792
QY	7298	CAGCTCCGGAGACGGTCAACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGT	7357
Db	8793	CAGCTCCGGAGACGGTCAACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGT	8852
QY	7358	CAGGCGCGTCAAGCGGTGTTGCGGGGTGTCGGGTGCTTAACTATCGGCGATCAGAG	7417
Db	8853	CAGGCGCGTCAAGCGGTGTTGCGGGGTGTCGGGTGCTTAACTATCGGCGATCAGAG	8912
QY	7418	CAGATTGTACTGAGATGACCAATATCGGTGTGAAATACCGCACAGATGCGTAAAGGAGA	7477
Db	8913	CAGATTGTACTGAGATGACCAATATCGGTGTGAAATACCGCACAGATGCGTAAAGGAGA	8972
QY	7478	AAATACCGCATCAGGCGCCATTCCGCAATTCAGGCTCGCAACTGTTTGGGAAGGGCGATCG	7537
Db	8973	AAATACCGCATCAGGCGCCATTCCGCAATTCAGGCTCGCAACTGTTTGGGAAGGGCGATCG	9032
QY	7538	GTGCGGCGCTTTCGCTATTATACGCCAGCTGGCGAAAAAGGGGGATGTGCTGCAAGCGCATTA	7597
Db	9033	GTGCGGCGCTTTCGCTATTATACGCCAGCTGGCGAAAAAGGGGGATGTGCTGCAAGCGCATTA	9092
QY	7598	AGTTGGGTAAACCGCAGGGTTTTCCAGTACGACGATGTGTAACACGACGGCAGTGCC	7654
Db	9093	AGTTGGGTAAACCGCAGGGTTTTCCAGTACGACGATGTGTAACACGACGGCAGTGCC	9149

Search completed: May 25, 2005, 19:58:51  
Job time : 1017 secs

**THIS PAGE BLANK (USPTO)**

Result No.	Score	Query Match	Length	DB	ID	Description
1	7654	100.0	7654	13	US-10-006-773-1	Sequence 1, Appl
2	7654	100.0	7654	13	US-10-006-771A-1	Sequence 1, Appl
3	3900	51.0	6963	19	US-10-913-288-1	Sequence 1, Appl
4	3359.8	43.9	7277	19	US-10-789-938B-3	Sequence 3, Appl
5	3354.4	43.8	7295	19	US-10-789-938B-2	Sequence 2, Appl
6	3349	43.8	9320	18	US-10-471-065-20	Sequence 20, Appl
7	3213	42.0	5782	10	US-09-982-223A-1	Sequence 1, Appl
8	3151.6	41.2	6254	16	US-10-317-078-1	Sequence 1, Appl
9	3151.6	41.2	6254	19	US-10-789-938B-1	Sequence 1, Appl
10	3103.4	40.5	6444	9	US-09-808-743-2	Sequence 2, Appl
11	3082.2	40.3	6219	9	US-09-963-206B-4	Sequence 4, Appl

Qy 1 AAGTTTGATCCTTCGAGGTCGACTCTAGGCACATAAAGAAAAACATAAACAAGCT 60

Dp 1 AAGTTTGATCCTTCGAGGTCGACTCTAGGCACATAAAGAAAAACATAAACAAGCT 60

QY 61 GCAGCCGACAGTGAAGAAACCGTTAAACCGTTGTTTAAATAAACTGAATATTTT 120  
 DB |||||  
 DB 61 GCAGCCGACAGTGAAGAAACCGTTAAACCGTTGTTTAAATAAACTGAATATTTT 120  
 QY 121 AGAGTCATTCTTTGTGTAGAAAGTACATGCGACGTAAAGAGGCCAAAGCAATCTGTG 180  
 DB |||||  
 DB 121 AGAGTCATTCTTTGTGTAGAAAGTACATGCGACGTAAAGAGGCCAAAGCAATCTGTG 180  
 QY 181 GAAAGCCGAGGCTGGAGGCCAGCAGTTTGATCCCTCTCGCGTGTACCTAAGGGTTT 240  
 DB |||||  
 DB 181 GAAAGCCGAGGCTGGAGGCCAGCAGTTTGATCCCTCTCGCGTGTACCTAAGGGTTT 240  
 QY 241 CTTAAATTTGTGTGTTTCTAAATCTTCCAGAGGTTTGTCTCATCTCACTTCCACTTCGGTG 300  
 DB |||||  
 DB 241 CTTAAATTTGTGTGTTTCTAAATCTTCCAGAGGTTTGTCTCATCTCACTTCCACTTCGGTG 300  
 QY 301 CACAATPACTTGGACGCGGATTTATCTGCTTAGCATCTATCGGTGGCCCTTCGATTTAGAGC 360  
 DB |||||  
 DB 301 CACAATPACTTGGACGCGGATTTACTGTCTTAGCATCTATCGGTGGCCCTTCGATTTAGAGC 360  
 QY 361 TGAACCTGAGGCCACTTCTGAGCTTGTAGGAGACACAAGACACGAGAGGCTGA 420  
 DB |||||  
 DB 361 TGAACCTGAGGCCACTTCTGAGCTTGTAAAGAGACACAAGACACGAGAGGCTGA 420  
 QY 421 CCGCGCAGACCTGTGGGCATTTTAAACAGGGCTCTCTGSETCTGTGGGAGGCAGGCTTA 480  
 DB |||||  
 DB 421 CCGCGCAGACCTGTGGGCATTTTAAACAGGGCTCTCTGSETCTGTGGGAGGCAGGCTTA 480  
 QY 481 CATPAGGTGCAATTAAGAAATATAATAATAAGCCCATATCAATTTGTGTCATCTTTTATTA 540  
 DB |||||  
 DB 481 CATPAGGTGCAATTAAGAAATATAATAATAAGCCCATATCAATTTGTGTCATCTTTTATTA 540  
 QY 541 AGCTCAAGTTTGAAGACCCCACTGTAGGTTTGGCAAGCTAGCTTAAAGTAAAGCCATT 600  
 DB |||||  
 DB 541 AGCTCAAGTTTGAAGACCCCACTGTAGGTTTGGCAAGCTAGCTTAAAGTAAAGCCATT 600  
 QY 601 TTGCAAGGATGGAATAACATAAATGAGAAATGAGAAAGTTCAGATCAAGAGTTAGGAACA 660  
 DB |||||  
 DB 601 TTGCAAGGATGGAATAACATAAATGAGAAATGAGAAAGTTCAGATCAAGAGTTAGGAACA 660  
 QY 661 GAGAGACAGCAATATGGGCCAAACAGGATATCTGTGGTAAAGCAGTTCTGCCCCGCTC 720  
 DB |||||  
 DB 661 GAGAGACAGCAATATGGGCCAAACAGGATATCTGTGGTAAAGCAGTTCTGCCCCGCTC 720  
 QY 721 AGGGCCCAAGAACAGTTGGAAACAGGAGAAATATGGGCCAAACAGGATATCTGTGGTAAAGCAG 780  
 DB |||||  
 DB 721 AGGGCCCAAGAACAGTTGGAAACAGGAGAAATATGGGCCAAACAGGATATCTGTGGTAAAGCAG 780  
 QY 781 TTCTGCCCCGGCTCAGGGCCCAAGAACAGATGTCCTCCAGATGCGGTCCCGCCCTCAGCA 840  
 DB |||||  
 DB 781 TTCTGCCCCGGCTCAGGGCCCAAGAACAGATGTCCTCCAGATGCGGTCCCGCCCTCAGCA 840  
 QY 841 GTTTCTAGAGAACATCAGATGTTTCCAGGGTGCCCCAAAGGACCTGAAATGACCCCTGTGC 900  
 DB |||||  
 DB 841 GTTTCTAGAGAACATCAGATGTTTCCAGGGTGCCCCAAAGGACCTGAAATGACCCCTGTGC 900  
 QY 901 CTTAATTTGAACCTAAGTTCGCTTCTGCTTCTGTTTCCGGGCTCTGCTCCCGC 960  
 DB |||||  
 DB 901 CTTAATTTGAACCTAAGTTCGCTTCTGCTTCTGTTTCCGGGCTCTGCTCCCGC 960  
 QY 961 AGCTCAATTAAGAGCCCAACCCCTCACTCGGCGCCAGTCCCTCCGATAGACTGCGT 1020  
 DB |||||  
 DB 961 AGCTCAATTAAGAGCCCAACCCCTCACTCGGCGCCAGTCCCTCCGATAGACTGCGT 1020  
 QY 1021 CGCCCGGGTACCCGTAATCCCAATAAAGCTCTTGTGTTTGTGATCCGAATCGTGACTC 1080  
 DB |||||  
 DB 1021 CGCCCGGGTACCCGTAATCCCAATAAAGCTCTTGTGTTTGTGATCCGAATCGTGACTC 1080  
 QY 1081 GCTGATCTTGGAGGGTCTCTCAGATTGATTTGACTGCCACCTCGGGGGTCTTTCATT 1140  
 DB |||||  
 DB 1081 GCTGATCTTGGAGGGTCTCTCCTCAGATTGATTTGACTGCCACCTCGGGGGTCTTTCATT 1140

QY 1141 TGGAGGTTCCACCGAGATTTTGGAGACCCCTGCTCCAGGACCAACGACCCCGCGCGGA 1200  
 DB |||||  
 DB 1141 TGGAGGTTCCACCGAGATTTTGGAGACCCCTGCTCCAGGACCAACGACCCCGCGCGGA 1200  
 QY 1201 GGTAAAGTGGCCAGCAACTTATCTGTGCTGTCTGTCGATTTCTAGTGTCTATGACTGATTT 1260  
 DB |||||  
 DB 1201 GGTAAAGTGGCCAGCAACTTATCTGTGCTGTCTGTCGATTTCTAGTGTCTATGACTGATTT 1260  
 QY 1261 TATGGCCCTGCGTCCGTACTAGTTAGCTAACTAGCTCTGTATCTGCGCGACCCGTGTGG 1320  
 DB |||||  
 DB 1261 TATGGCCCTGCGTCCGTACTAGTTAGCTAACTAGCTCTGTATCTGCGCGACCCGTGTGG 1320  
 QY 1321 AACTGACGAGTTTCGAAACACCGCGCGCAACCCCTGGAGACGCTCCAGGACCTTCGGGG 1380  
 DB |||||  
 DB 1321 AACTGACGAGTTTCGAAACACCGCGCGCAACCCCTGGAGACGCTCCAGGACCTTCGGGG 1380  
 QY 1381 CCGTTTTTGTGGCCCGACCTGAGTCTTAAATCCGATCGTTTTAGGACTCTTTTGTGTGAC 1440  
 DB |||||  
 DB 1381 CCGTTTTTGTGGCCCGACCTGAGTCTTAAATCCGATCGTTTTAGGACTCTTTTGTGTGAC 1440  
 QY 1441 CCCCCTTAGAGAGGATATGTGTTCTGTTAGGAGACGAACTTAAACAGTTCCTCCG 1500  
 DB |||||  
 DB 1441 CCCCCTTAGAGAGGATATGTGTTCTGTTAGGAGACGAACTTAAACAGTTCCTCCG 1500  
 QY 1501 CTCCCTCTGAATTTTGTGTTTTCGGTTCGGAACCGAGCGCGCGCGCTCTTGTCTGCT 1560  
 DB |||||  
 DB 1501 CTCCCTCTGAATTTTGTGTTTTCGGTTCGGAACCGAGCGCGCGCGCTCTTGTCTGCT 1560  
 QY 1561 GCAGCATCGTCTGTGTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620  
 DB |||||  
 DB 1561 GCAGCATCGTCTGTGTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620  
 QY 1621 GCCCGGCTAGACTCTGTTACCACTCCCTTAAAGTTTACCTTACCTTACCTTACCTTACCT 1680  
 DB |||||  
 DB 1621 GCCCGGCTAGACTCTGTTACCACTCCCTTAAAGTTTACCTTACCTTACCTTACCTTACCT 1680  
 QY 1681 AGCGATCGCTCACCAACAGTCGCTAGATGTCAAGAGAGAGAGTTCGGTGTACCTTCTGCT 1740  
 DB |||||  
 DB 1681 AGCGATCGCTCACCAACAGTCGCTAGATGTCAAGAGAGAGAGTTCGGTGTACCTTCTGCT 1740  
 QY 1741 CTGCAAGATGGCCAACTTTTAAAGTTCGCTAGATGTCCGCGAGAGAGAGTTCGAGAGCC 1800  
 DB |||||  
 DB 1741 CTGCAAGATGGCCAACTTTTAAAGTTCGCTAGATGTCCGCGAGAGAGAGTTCGAGAGCC 1800  
 QY 1801 TCATCACCAGGTTAAGATCAAGGTCCTTTTACCTTGGCCCGCATGGAACCCAGAGCAG 1860  
 DB |||||  
 DB 1801 TCATCACCAGGTTAAGATCAAGGTCCTTTTACCTTGGCCCGCATGGAACCCAGAGCAG 1860  
 QY 1861 TCCCTTACATCTGACCTGGGAAGCTTTGGCTTTTGAACCCCTCTCCCTGGGTCAAGCCCT 1920  
 DB |||||  
 DB 1861 TCCCTTACATCTGACCTGGGAAGCTTTGGCTTTTGAACCCCTCTCCCTGGGTCAAGCCCT 1920  
 QY 1921 TTGTACAACCTTAAGCTCCGCTCTCTTCTCCATCCGCGCCGCTCTCTCCGCTTGAAC 1980  
 DB |||||  
 DB 1921 TTGTACAACCTTAAGCTCCGCTCTCTTCTCCATCCGCGCCGCTCTCTCCGCTTGAAC 1980  
 QY 1981 CTCCTCTGTCAGCCCGCTCGATCTCTCTTTATCCAGCCCTCTCTCTCTCTAGGG 2040  
 DB |||||  
 DB 1981 CTCCTCTGTCAGCCCGCTCGATCTCTCTTTATCCAGCCCTCTCTCTCTCTAGGG 2040  
 QY 2041 CCCCCTATGGCCATATGAGATCTTATATGGGCGACCCCGCCCTCTGTAAACTTCCCTG 2100  
 DB |||||  
 DB 2041 CCCCCTATGGCCATATGAGATCTTATATGGGCGACCCCGCCCTCTGTAAACTTCCCTG 2100  
 QY 2101 ACCCTGACATGAAGAGTTACTAAACAGCCCTCTCTCTCAAGCTCACTTACAGGCTTCTA 2160  
 DB |||||  
 DB 2101 ACCCTGACATGAAGAGTTACTAAACAGCCCTCTCTCTCAAGCTCACTTACAGGCTTCTA 2160  
 QY 2161 CTTAGTCAGACAGAGTCTGGAGACCTCTGGCGGAGCCTTACAGACACTTGGACCG 2220  
 DB |||||  
 DB 2161 CTTAGTCAGACAGAGTCTGGAGACCTCTGGCGGAGCCTTACAGACAACTGGACCG 2220  
 QY 2221 ACCGGTGGTACCTCACCTTACCGAGTCGGCGGACACAGTGTGGTCCGCGCACACAGAC 2280

|||||  
2221 ACCGGTGGTACCTCAACCTTACCGAGTCGGCCACACAGTGTGGGTCCGCCACACAGAC 2280  
QY  
2281 TAAGAACCTTAGAACCTCGCTGGAAGGACCTTTACACAGTCTCTGTGACACACCCACCCG 2340  
Db  
2281 TAAGAACCTTAGAACCTCGCTGGAAGGACCTTTACACAGTCTCTGTGACACACCCACCCG 2340  
QY  
2341 CCTCAAGTAGACGGCATCGAGCTTGGATACACCGCGCCACGCTGAAGGCTGCGACCC 2400  
Db  
2341 CCTCAAGTAGACGGCATCGAGCTTGGATACACCGCGCCACGCTGAAGGCTGCGACCC 2400  
QY  
2401 CGGGGTGGACCATCCTCTAGACTGCCATGGATGGAGTGTATCATCTCTTTTGGTA 2460  
Db  
2401 CGGGGTGGACCATCCTCTAGACTGCCATGGATGGAGTGTATCATCTCTTTTGGTA 2460  
QY  
2461 GCAACAGCTACAGGTGTCCACTCCGACATCCAGCTGACCCAGAGCCCAAGCAGCTGAGC 2520  
Db  
2461 GCAACAGCTACAGGTGTCCACTCCGACATCCAGCTGACCCAGAGCCCAAGCAGCTGAGC 2520  
QY  
2521 GCAGCGTGGTGAACAGATGACCATCACTGTAAAGCCAGTCAAGATGTGGGTACTTCT 2580  
Db  
2521 GCAGCGTGGTGAACAGATGACCATCACTGTAAAGCCAGTCAAGATGTGGGTACTTCT 2580  
QY  
2581 GTAGCTTGGTACCAGAGCCAGGTAAAGCTCCAAAGCTGCTGATCTACTGGACATCC 2640  
Db  
2581 GTAGCTTGGTACCAGAGCCAGGTAAAGCTCCAAAGCTGCTGATCTACTGGACATCC 2640  
QY  
2641 ACCCGCACACTGGTGTCCAAAGCAGATTACGCGGTAGCGGTACCGACTTCACC 2700  
Db  
2641 ACCCGCACACTGGTGTCCAAAGCAGATTACGCGGTAGCGGTACCGACTTCACC 2700  
QY  
2701 TTCAACCATCAGCAGCTCCAGCCAGAGGACATCGCCACTACTACTGCCAGCAATATAGC 2760  
Db  
2701 TTCAACCATCAGCAGCTCCAGCCAGAGGACATCGCCACTACTACTGCCAGCAATATAGC 2760  
QY  
2761 CTCTATCGGTGTTCGGCCAGAGGACCAAGGTGGAAATCAAACGAGGTGGCTCAGGATCG 2820  
Db  
2761 CTCTATCGGTGTTCGGCCAGAGGACCAAGGTGGAAATCAAACGAGGTGGCTCAGGATCG 2820  
QY  
2821 GGTGATCCGGCTCTGGTGGCTCAGATCGGAGGTCCAACTGGTGAGAGCGGTGGAGGT 2880  
Db  
2821 GGTGATCCGGCTCTGGTGGCTCAGATCGGAGGTCCAACTGGTGAGAGCGGTGGAGGT 2880  
QY  
2881 GTTGTCAACCTGGCGGTCTCGCTGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 2940  
Db  
2881 GTTGTCAACCTGGCGGTCTCGCTGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 2940  
QY  
2941 ACATATTCGATGAGTTGGGTGAGACAGGCACTTGAAAAGGTCTTGGATGGAGAA 3000  
Db  
2941 ACATATTCGATGAGTTGGGTGAGACAGGCACTTGAAAAGGTCTTGGATGGAGAA 3000  
QY  
3001 ATTCAATCCAGATAGCAGTACGATTAACCTATGCGCGCTCTTAAGGATAGATTTACATA 3060  
Db  
3001 ATTCAATCCAGATAGCAGTACGATTAACCTATGCGCGCTCTTAAGGATAGATTTACATA 3060  
QY  
3061 TCGGAGACAAACCCGAGCAACATTTGTTCTGCAATTTGGACAGCTCAGACCCGAGAC 3120  
Db  
3061 TCGGAGACAAACCCGAGCAACATTTGTTCTGCAATTTGGACAGCTCAGACCCGAGAC 3120  
QY  
3121 ACCGGGTCTATTTTGTGCAAGCTTTACTTGGCTTTCCCTGCTGCTGCTGCTGCTGCTGCTG 3180  
Db  
3121 ACCGGGTCTATTTTGTGCAAGCTTTACTTGGCTTTCCCTGCTGCTGCTGCTGCTGCTGCTG 3180  
QY  
3181 CAAGGACCCCGGTCAACCTCTCAGTCTTAAGCCACACAGCCAGCCAGCCCGGACCA 3240  
Db  
3181 CAAGGACCCCGGTCAACCTCTCAGTCTTAAGCCACACAGCCAGCCAGCCCGGACCA 3240  
QY  
3241 CCAACACCGGCGCCACCATCGCTCGAGCCCTGCTCCCTGCGCCAGAGCGCGCTCGG 3300  
Db  
3241 CCAACACCGGCGCCACCATCGCTCGAGCCCTGCTCCCTGCGCCAGAGCGCGCTCGG 3300  
QY  
3301 CCAGCGCGGGGGCGGAGTGACACAGAGGGGTGAGATTCGCGCTCGATCCCAAACTC 3360  
|||||

Db  
3301 CCAGCGCGGGGGCGGAGTGACACAGAGGGGGCTGGACTTCGCCCTCGATCCCAAACTC 3360  
QY  
3361 TGCTACCTCTGGATGGAATCCTCTTCACTATATGTGTCACTTCTCACTGCTGTCTCTG 3420  
Db  
3361 TGCTACCTCTGGATGGAATCCTCTTCACTATATGTGTCACTTCTCACTGCTGTCTCTG 3420  
QY  
3421 AGAGTGAAGTTTCAGAGGAGCGAGAGCCCGCGGTACAGAGGGGCGAGAACGAGCTC 3480  
Db  
3421 AGAGTGAAGTTTCAGAGGAGCGAGAGCCCGCGGTACAGAGGGGCGAGAACGAGCTC 3480  
QY  
3481 TATAACGAGCTCAATCTTAGGACGAAGAGAGGAGTACGATGTTTTGGACAAGAGAGCTGGC 3540  
Db  
3481 TATAACGAGCTCAATCTTAGGACGAAGAGAGGAGTACGATGTTTTGGACAAGAGAGCTGGC 3540  
QY  
3541 CGGACCTCTAGATGCGGGGAAAGCCAGAGAGGAGAAACCTCTCAGAGAGGCTGTACAAT 3600  
Db  
3541 CGGACCTCTAGATGCGGGGAAAGCCAGAGAGGAGAAACCTCTCAGAGAGGCTGTACAAT 3600  
QY  
3601 GAATCTGCAGAAAGATTAAGATGGCGAGGCTTACAGTGAGATTTGGATGAAAGGCGAGCGC 3660  
Db  
3601 GAATCTGCAGAAAGATTAAGATGGCGAGGCTTACAGTGAGATTTGGATGAAAGGCGAGCGC 3660  
QY  
3661 CGGAGGGGCAAGGGGCGAGTGGCTTTTACAGGGTCTCAGTACAGCCACCAAGAGACACC 3720  
Db  
3661 CGGAGGGGCAAGGGGCGAGTGGCTTTTACAGGGTCTCAGTACAGCCACCAAGAGACACC 3720  
QY  
3721 TACGACGCTTACATGTCAGGCTCGCCCTCGCTAACTCGAGCGCGCGCGGATCCG 3780  
Db  
3721 TACGACGCTTACATGTCAGGCTCGCCCTCGCTAACTCGAGCGCGCGCGGATCCG 3780  
QY  
3781 GATTAGTCCAATTTGTTAAAGACAGGATATCAGTGGTTCAGGCTCTAGTTTGACTCAAC 3840  
Db  
3781 GATTAGTCCAATTTGTTAAAGACAGGATATCAGTGGTTCAGGCTCTAGTTTGACTCAAC 3840  
QY  
3841 AATATCACAGCTGAAGCTTATAGATACGAGCCATAGATAAATAAAGATTTTATTTA 3900  
Db  
3841 AATATCACAGCTGAAGCTTATAGATACGAGCCATAGATAAATAAAGATTTTATTTA 3900  
QY  
3901 GTCTCCAGAAAAGGGGGAAATGAAGACCCACTGTAGGTTTGGCAAGCTAGCTTAAG 3960  
Db  
3901 GTCTCCAGAAAAGGGGGAAATGAAGACCCACTGTAGGTTTGGCAAGCTAGCTTAAG 3960  
QY  
3961 TAACGCCATTTTGAAGGCGATGGAATAATACATACTGAGAAATAGAGAGTTTCAAGTCAAG 4020  
Db  
3961 TAACGCCATTTTGAAGGCGATGGAATAATACATACTGAGAAATAGAGAGTTTCAAGTCAAG 4020  
QY  
4021 GTTAGGAAACAGAGAGACAGAGAAATATGGGCCAAACAGGATATCTGTGGTAAGCATTC 4080  
Db  
4021 GTTAGGAAACAGAGAGACAGAGAAATATGGGCCAAACAGGATATCTGTGGTAAGCATTC 4080  
QY  
4081 TGCCCGCTCAGGGCCAGAGACAGTGGAAACAGAGAAATATGGGCCAAACAGGATATCTG 4140  
Db  
4081 TGCCCGCTCAGGGCCAGAGACAGTGGAAACAGAGAAATATGGGCCAAACAGGATATCTG 4140  
QY  
4141 TGGTAAGCAGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4200  
Db  
4141 TGGTAAGCAGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4200  
QY  
4201 GCGCTCAGCAGTTCTTAGAGAACCATCAGATGTTTTCAAGGTCGCCCAAGGACCTGAAAT 4260  
Db  
4201 GCGCTCAGCAGTTCTTAGAGAACCATCAGATGTTTTCAAGGTCGCCCAAGGACCTGAAAT 4260  
QY  
4261 GACCTGTGCTTATTTGAACTAACCAATCAGTTCGCTTCTCGTTCGTCGCGCGCTT 4320  
Db  
4261 GACCTGTGCTTATTTGAACTAACCAATCAGTTCGCTTCTCGTTCGTCGCGCGCTT 4320  
QY  
4321 CTGCTCCCGAGCTCAATATAAGAGCCCAACCCCTCACTCGCGCGCCAGTCTCTCGA 4380  
Db  
4321 CTGCTCCCGAGCTCAATATAAGAGCCCAACCCCTCACTCGCGCGCCAGTCTCTCGA 4380  
QY  
4381 TAGACTCGCTCGCGGTACCCGTTCTCAATATAACCCCTCTTTCAGTTCGATCCGACT 4440  
Db  
4381 TAGACTCGCTCGCGGTACCCGTTCTCAATATAACCCCTCTTTCAGTTCGATCCGACT 4440



QY 4441 CGTGGTCTCGCTGTTCTTTGGGAGGGTCTCTCTGAGTGATTTGACTACCGGTGAGGGGT 4500  
 Db 4441 CGTGGTCTCGCTGTTCTTTGGGAGGGTCTCTCTGAGTGATTTGACTACCGGTGAGGGGT 4500  
 QY 4501 CTTTCAGTTTCTCCCACTTACACAGGTCTCACTAAACATTCCTGATGTGCGCAGGACATC 4560  
 Db 4501 CTTTCAGTTTCTCCCACTTACACAGGTCTCACTAAACATTCCTGATGTGCGCAGGACATC 4560  
 QY 4561 CGTCAGCCCGGTTTTGTTTATAATAAATGCAAGACAGTGTTCCTTCAAGCCAGACT 4620  
 Db 4561 CGTCAGCCCGGTTTTGTTTATAATAAATGCAAGACAGTGTTCCTTCAAGCCAGACT 4620  
 QY 4621 ACATCTGACTCTCGGCTTTATAAAGAAATGTTGAAGGCTCTGTGGAATATCTGCAACA 4680  
 Db 4621 ACATCTGACTCTCGGCTTTATAAAGAAATGTTGAAGGCTCTGTGGAATATCTGCAACA 4680  
 QY 4681 CGACTTTTTAAGATTTTTATGCTCTCGGATGAGGATTTAGTCAATCTATCTCGTCTA 4740  
 Db 4681 CGACTTTTTAAGATTTTTATGCTCTCGGATGAGGATTTAGTCAATCTATCTCGTCTA 4740  
 QY 4741 TTTTGTGCTTCTCGTATTTTAAATTTCTAGTTTGGCACTCCCTTCTTGAGAGCACGGC 4800  
 Db 4741 TTTTGTGCTTCTCGTATTTTAAATTTCTAGTTTGGCACTCCCTTCTTGAGAGCACGGC 4800  
 QY 4801 GATTGACAGAGTAAATTAATCTCTGAGGCGAGGCTCTGTGAAAAGGTTGCTTGGGCTCAG 4860  
 Db 4801 GATTGACAGAGTAAATTAATCTCTGAGGCGAGGCTCTGTGAAAAGGTTGCTTGGGCTCAG 4860  
 QY 4861 TGTGAGATTTTGCATAAAGAGGGTCTGCGCTGTGTACAGACAGATCGGAATCTAGA 4920  
 Db 4861 TGTGAGATTTTGCATAAAGAGGGTCTGCGCTGTGTACAGACAGATCGGAATCTAGA 4920  
 QY 4921 GTGCATACACAGTCCCGGGTTCGGGGTCTGATCTCAGGGCATCTTTGCGCTAGAG 4980  
 Db 4921 GTGCATACACAGTCCCGGGTTCGGGGTCTGATCTCAGGGCATCTTTGCGCTAGAG 4980  
 QY 4981 ATCTCTACCGCGGACGATCGTGGCGGTACCGAGCTCGAATTCGTAATCATGTCTAT 5040  
 Db 4981 ATCTCTACCGCGGACGATCGTGGCGGTACCGAGCTCGAATTCGTAATCATGTCTAT 5040  
 QY 5041 AGCTGTTTCTGTGTGAAATTTGTTATCGCTCACAAATCCACACATACGAGCGGAA 5100  
 Db 5041 AGCTGTTTCTGTGTGAAATTTGTTATCGCTCACAAATCCACACATACGAGCGGAA 5100  
 QY 5101 GCATAAAGTGTAAAGCTTGGGGTGCCTAATGAGTGAGCTAACTCAATTAATTTGGCTTGC 5160  
 Db 5101 GCATAAAGTGTAAAGCTTGGGGTGCCTAATGAGTGAGCTAACTCAATTAATTTGGCTTGC 5160  
 QY 5161 GCTCACTGCGCTTTTCCAGTCCGGAAACCTGTGCGCAGCTGCATTAATGAATCGGOC 5220  
 Db 5161 GCTCACTGCGCTTTTCCAGTCCGGAAACCTGTGCGCAGCTGCATTAATGAATCGGOC 5220  
 QY 5221 AACGCGCGGGAGAGCGGTTTTCGTAATTTGGGCGCTCTTTCCGTTTCTCGCTCACTGACT 5280  
 Db 5221 AACGCGCGGGAGAGCGGTTTTCGTAATTTGGGCGCTCTTTCCGTTTCTCGCTCACTGACT 5280  
 QY 5281 CGTGGCTCGGTCGTTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAAGCGGTAAATAC 5340  
 Db 5281 CGTGGCTCGGTCGTTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAAGCGGTAAATAC 5340  
 QY 5341 GGTATCCACAGAAATCAGGGGATTAACGAGGAAAGAAATGATGAGCAAAAGGCCAGCAA 5400  
 Db 5341 GGTATCCACAGAAATCAGGGGATTAACGAGGAAAGAAATGATGAGCAAAAGGCCAGCAA 5400  
 QY 5401 AGGCCAGGAACCGTAAAAAGGCGCGTTCGTCGGCTTTTTCATAGGCTTCGCGCCCGCTG 5460  
 Db 5401 AGGCCAGGAACCGTAAAAAGGCGCGTTCGTCGGCTTTTTCATAGGCTTCGCGCCCGCTG 5460  
 QY 5461 ACAGAGCATCAAAAAATCAGCGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAA 5520  
 Db 5461 ACAGAGCATCAAAAAATCAGCGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAA 5520

QY 5521 GATACACGCGGTTCCCGCTCGAAGCTCCCTGCTGGCTCTCTCTGTTCCGACCCCTGCGCG 5580  
 Db 5521 GATACACGCGGTTCCCGCTCGAAGCTCCCTGCTGGCTCTCTCTGTTCCGACCCCTGCGCG 5580  
 QY 5581 TTACCGGATACCTGTCCGCTTTCTCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAC 5640  
 Db 5581 TTACCGGATACCTGTCCGCTTTCTCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAC 5640  
 QY 5641 GCTGTAGGTATCTCAGTTTCGGTGTAGGTCTGCTCCAAAGCTGGCTGTGTGTCACGAAAC 5700  
 Db 5641 GCTGTAGGTATCTCAGTTTCGGTGTAGGTCTGCTCCAAAGCTGGCTGTGTGTCACGAAAC 5700  
 QY 5701 CCCCCGTTTCAGCCCCGACCGCTGCGCTTATCCGGTAACTATCTGCTTTCAGTCTCAACCCGG 5760  
 Db 5701 CCCCCGTTTCAGCCCCGACCGCTGCGCTTATCCGGTAACTATCTGCTTTCAGTCTCAACCCGG 5760  
 QY 5761 TAAAGACAGACTTATTCGCACTGCGAGCAGCCACTGGTAAACAGGATTTAGCAGAGCGAGT 5820  
 Db 5761 TAAAGACAGACTTATTCGCACTGCGAGCAGCCACTGGTAAACAGGATTTAGCAGAGCGAGT 5820  
 QY 5821 ATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACTACCGGCTACACTAGAAGA 5880  
 Db 5821 ATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACTACCGGCTACACTAGAAGA 5880  
 QY 5881 CAGTATTTGGTATCTGCGCTCTGCTGAAAGCCAGTTTACCTTCGGAAGAAAGAGTTGGTAGCT 5940  
 Db 5881 CAGTATTTGGTATCTGCGCTCTGCTGAAAGCCAGTTTACCTTCGGAAGAAAGAGTTGGTAGCT 5940  
 QY 5941 CTTGATTCGGGCAAAACAAACCCAGCTGGTAGCGGTGGTTTTTTTGTGTAAGACGACAGA 6000  
 Db 5941 CTTGATTCGGGCAAAACAAACCCAGCTGGTAGCGGTGGTTTTTTTGTGTAAGACGACAGA 6000  
 QY 6001 TTACGCGCAGAAAAAAGGATCTCAAGAGATCCCTTTGATCTTTTCTACGGGCTCTGACG 6060  
 Db 6001 TTACGCGCAGAAAAAAGGATCTCAAGAGATCCCTTTGATCTTTTCTACGGGCTCTGACG 6060  
 QY 6061 CTCAGTGAACGAAACCTCACTTAAAGGATTTTGGTCAAGAGATTTCAAAAAGGATCT 6120  
 Db 6061 CTCAGTGAACGAAACCTCACTTAAAGGATTTTGGTCAAGAGATTTCAAAAAGGATCT 6120  
 QY 6121 TCACCTAGATCTTTTAAATTTAAATGAAAGTTTAAATCAATCTAAAGTATATATAGT 6180  
 Db 6121 TCACCTAGATCTTTTAAATTTAAATGAAAGTTTAAATCAATCTAAAGTATATATAGT 6180  
 QY 6181 AAATTTGCTGTGACAGTTTACCAATCTTAAATCAGTGAGGACCTATCTCAGCGATCTGTC 6240  
 Db 6181 AAATTTGCTGTGACAGTTTACCAATCTTAAATCAGTGAGGACCTATCTCAGCGATCTGTC 6240  
 QY 6241 TATTTTCGTTTCATCCATAGTTGCTGACTCCCGCTCGTGTAGATACTACGATACGGGAGG 6300  
 Db 6241 TATTTTCGTTTCATCCATAGTTGCTGACTCCCGCTCGTGTAGATACTACGATACGGGAGG 6300  
 QY 6301 GCTTACCATCTGGCCCCAGTGTCAATGATACCGCGAGACCCACGCTCACCGGCTCCAG 6360  
 Db 6301 GCTTACCATCTGGCCCCAGTGTCAATGATACCGCGAGACCCACGCTCACCGGCTCCAG 6360  
 QY 6361 ATTTATCAGCAATAAACACGCGGAGGCGGAGGAGAGTGTCTGCAACTT 6420  
 Db 6361 ATTTATCAGCAATAAACACGCGGAGGCGGAGGAGTGTCTGCAACTT 6420  
 QY 6421 TATCCGCTCCATCCAGTCTTATTAATTTGTCGGGAAAGCTAGAGTAAGTAGTTCCGCGAG 6480  
 Db 6421 TATCCGCTCCATCCAGTCTTATTAATTTGTCGGGAAAGCTAGAGTAAGTAGTTCCGCGAG 6480  
 QY 6481 TTAATAGTTTGGCAACCGTTTGCATTTGCTACAGGCTCGTGGTGTCAACGCTCGTGT 6540  
 Db 6481 TTAATAGTTTGGCAACCGTTTGCATTTGCTACAGGCTCGTGGTGTCAACGCTCGTGT 6540  
 QY 6541 TGGTATGGCTTCAATTCAGCTCCGTTCCCAAGTCAAGGCGAGTTACATGATCCCCCAT 6600  
 Db 6541 TGGTATGGCTTCAATTCAGCTCCGTTCCCAAGTCAAGGCGAGTTACATGATCCCCCAT 6600  
 QY 6601 GTTGTGCAAAAAAGCGGTTAGCTCTCTCGTCTCCGATCGTTGTGTCAGAAAGTAAGTTGCG 6660

```

6601 GTTGTGCAAAAAGCGTTAGCTCTTCGGTCTCCGATCGTTGTGACAGTAAAGTTGGC 6660
6661 CGCAGTGTATCACTCATGTTATGGCAGCACTGCATATAATCTCTTCTTCTGTCATGCCATC 6720
6661 CGCAGTGTATCACTCATGTTATGGCAGCACTGCATATAATCTCTTCTTCTGTCATGCCATC 6720
6721 CGTAAAGTCTTTCTGTGACTGGTGAAGTACTCAACCAAGTCAATCTCTGAGAAATAGTGAT 6780
6721 CGTAAAGTCTTTCTGTGACTGGTGAAGTACTCAACCAAGTCAATCTCTGAGAAATAGTGAT 6780
6781 GCGCGACCGAGTTGCTCTCTGCGCGCGCTCAATACGGGATAATACCGCGCCACATAGCAG 6840
6781 GCGCGACCGAGTTGCTCTCTGCGCGCGCTCAATACGGGATAATACCGCGCCACATAGCAG 6840
6841 AACTTTAAAGTCTCATCATTTGAAAAAGTTTTCGGGGGAAAACTCTCTCAAGGATCTT 6900
6841 AACTTTAAAGTCTCATCATTTGAAAAAGTTTTCGGGGGAAAACTCTCTCAAGGATCTT 6900
6901 ACCGCTGTGAGATCCAGTTCGATTAACCACTCGTGCACCACTGATCTTCAGCATC 6960
6901 ACCGCTGTGAGATCCAGTTCGATTAACCACTCGTGCACCACTGATCTTCAGCATC 6960
6961 TTTTACTTTTACCAGCGTTCTTGGGTGAGCAAAAACAGAAAGGCAAAATGCGCAAAAAA 7020
6961 TTTTACTTTTACCAGCGTTCTTGGGTGAGCAAAAACAGAAAGGCAAAATGCGCAAAAAA 7020
7021 GGGATAAGGCGCACACGGAATGTTGAATACTCATACTCTCTCTTTTCAATATTATTG 7080
7021 GGGATAAGGCGCACACGGAATGTTGAATACTCATACTCTCTCTTTTCAATATTATTG 7080
7081 AAGCAATTATCAGGGTTATTGCTCATGAGCGGATACATATTGAATGTTATTAGAAAAA 7140
7081 AAGCAATTATCAGGGTTATTGCTCATGAGCGGATACATATTGAATGTTATTAGAAAAA 7140
7141 TAAACAAATAGGGTTCCGCGCACATTTTCCCGAAAAGTGCACCTGACGCTCTAAGAAC 7200
7141 TAAACAAATAGGGTTCCGCGCACATTTTCCCGAAAAGTGCACCTGACGCTCTAAGAAC 7200
7201 CATTAATTATCATGACATTAACTATAAAATAGGCGTATCACAGGCGCTTCTGCTCGC 7260
7201 CATTAATTATCATGACATTAACTATAAAATAGGCGTATCACAGGCGCTTCTGCTCGC 7260
7261 CGCTTTCGGTGATGACGGTGAACCTCTGACACATGACGCTCCCGGAGACGGTCAACAGC 7320
7261 CGCTTTCGGTGATGACGGTGAACCTCTGACACATGACGCTCCCGGAGACGGTCAACAGC 7320
7321 TTGCTGTAAAGGATGCGGAGCAGACAAGCCGCTCAGGCGCGGTGAGGGGTGTTGG 7380
7321 TTGCTGTAAAGGATGCGGAGCAGACAAGCCGCTCAGGCGCGGTGAGGGGTGTTGG 7380
7381 CGGGTGTCCGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGACCA 7440
7381 CGGGTGTCCGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGACCA 7440
7441 TATCGGTGTGAAATACCGCAGATGCGTAAAGGAGAAATAACCGCATCAGCGCCATTTC 7500
7441 TATCGGTGTGAAATACCGCAGATGCGTAAAGGAGAAATAACCGCATCAGCGCCATTTC 7500
7501 GCCATTACGGCTGCGCAACTGTTGGGAAGGCGGATCGGTGCGGGCTCTTTCGCTATTACG 7560
7501 GCCATTACGGCTGCGCAACTGTTGGGAAGGCGGATCGGTGCGGGCTCTTTCGCTATTACG 7560
7561 CCAGTCGCAAAAGGGGATGCTGCAAGCGGATTAAAGTTGGGTAAAGCGGGTTTC 7620
7561 CCAGTCGCAAAAGGGGATGCTGCAAGCGGATTAAAGTTGGGTAAAGCGGGTTTC 7620
7621 CCAGTCACGAGCTTTGAAAAACGACGGCCAGTGCC 7654
7621 CCAGTCACGAGCTTTGAAAAACGACGGCCAGTGCC 7654
```

US-10-006-771A-1  
; Sequence 1, Application US/10006771A  
; Publication No. US20020165360A1  
; GENERAL INFORMATION:  
; APPLICANT: Junghans, Richard P.  
; TITLE OF INVENTION: Chimeric Effector Cell Receptors Against Carcinoembryonic Antigen  
; FILE REFERENCE: 002  
; CURRENT APPLICATION NUMBER: US/10/006,771A  
; CURRENT FILING DATE: 2002-06-04  
; PRIOR APPLICATION NUMBER: 60/250,090  
; PRIOR FILING DATE: 2000-11-30  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 7654  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (2428)..(3759)  
; OTHER INFORMATION: Chimeric IgTCR sequence contained in retroviral vector. Retrovi  
; OTHER INFORMATION: al vector sequence (non-coding regions) are incidental to the inv  
; OTHER INFORMATION: ention. The translated (coding region) is relevant to the invent  
; OTHER INFORMATION: ion. (pertinent to Figure 3.)

US-10-006-771A-1  
Query Match 100.0%; Score 7654; DB 13; Length 7654;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 7654; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AAGCTTGATGCTGCGAGTGCAGTCTAGGCACATATAAGAAAAACATTAACCAAGCT 60  
Db 1 AAGCTTGATGCTGCGAGTGCAGTCTAGGCACATATAAGAAAAACATTAACCAAGCT 60  
Qy 61 GCACCCGAGCAGTGAAGAAGCCGTTAAACCGTTTGTGTTTAAATAAATGAAATTTT 120  
Db 61 GCACCCGAGCAGTGAAGAAGCCGTTAAACCGTTTGTGTTTAAATAAATGAAATTTT 120  
Qy 121 AGAGTCATTTCTTTGGTAGAAAGTACATTTGGCAGCGTAAAGAGGCCAAAGCAATCTGTG 180  
Db 121 AGAGTCATTTCTTTGGTAGAAAGTACATTTGGCAGCGTAAAGAGGCCAAAGCAATCTGTG 180  
Qy 181 GAAAGCCAGCTGGGAGCCAGCAGTTTGATCCCTCTCTGGCGTGTACTAAGGGTTT 240  
Db 181 GAAAGCCAGCTGGGAGCCAGCAGTTTGATCCCTCTCTGGCGTGTACTAAGGGTTT 240  
Qy 241 CTTAAATTTGTGGTTTCTAAATCTTCCAGAGGGTTTGTCTCAATTCACCTCCAGTCCGTTG 300  
Db 241 CTTAAATTTGTGGTTTCTAAATCTTCCAGAGGGTTTGTCTCAATTCACCTCCAGTCCGTTG 300  
Qy 301 CACAATCTTTGGAGCCGGATTTACTGTCTTAGCATCTATCGGTGGCCCTTCGATTGAGGC 360  
Db 301 CACAATCTTTGGAGCCGGATTTACTGTCTTAGCATCTATCGGTGGCCCTTCGATTGAGGC 360  
Qy 361 TGAACTTGAGGCCCACTTCTTACGCTTTGTTAAGGAGAGCACAAGCACCAGAGAGGCTGA 420  
Db 361 TGAACTTGAGGCCCACTTCTTACGCTTTGTTAAGGAGAGCACAAGCACCAGAGAGGCTGA 420  
Qy 421 CCGGCGAGACCTGTGGGCATTTTAAACAAGGCCCTCTGGTCTGTGGAGGCGAGGCTTA 480  
Db 421 CCGGCGAGACCTGTGGGCATTTTAAACAAGGCCCTCTGGTCTGTGGAGGCGAGGCTTA 480  
Qy 481 CATAAAGTGCAAATTAAGAAATATAAATAAAGCCCATATCAATTTGTTCATCTTTTTTA 540  
Db 481 CATAAAGTGCAAATTAAGAAATATAAATAAAGCCCATATCAATTTGTTCATCTTTTTTA 540  
Qy 541 AGCTCAAGTTTGAAGACCCCACTCTAGTGTTTGGCAAGCTAGCTTAAGTAAAGCCATT 600  
Db 541 AGCTCAAGTTTGAAGACCCCACTCTAGTGTTTGGCAAGCTAGCTTAAGTAAAGCCATT 600  
Qy 601 TTGCAAGGCATGGAATAATACATACTGAGAAATAGAGAGTTTCAGATCAAGTTTAGGAACA 660  
Db 601 TTGCAAGGCATGGAATAATACATACTGAGAAATAGAGAGTTTCAGATCAAGTTTAGGAACA 660





QY 5041 AGCTGTTTCTGTGTGAATTTGTTATCCGCTCACAATTCACAACAACATACGAGCGGAA 5100  
 DB |||||  
 QY 5041 AGCTGTTTCTGTGTGAATTTGTTATCCGCTCACAATTCACAACAACATACGAGCGGAA 5100  
 DB |||||  
 QY 5101 GCATAAAGTGTAAAGCTCGGGTGCTTAATAGTGTAGCTAACTCAATTAATTTGGGTTGC 5160  
 DB |||||  
 QY 5101 GCATAAAGTGTAAAGCTCGGGTGCTTAATAGTGTAGCTAACTCAATTAATTTGGGTTGC 5160  
 DB |||||  
 QY 5161 GCTCACTGCCGCTTTCCAGTCCGGGAAACCTGTGTGCCAGTGCATTAATGAATCGGCC 5220  
 DB |||||  
 QY 5161 GCTCACTGCCGCTTTCCAGTCCGGGAAACCTGTGTGCCAGTGCATTAATGAATCGGCC 5220  
 DB |||||  
 QY 5221 AACCGCGGGGAGAGCGGTTTGGGTATTGGCGCTCTTCCGCTTCCGCTCCTCGCTCACTGACT 5280  
 DB |||||  
 QY 5221 AACCGCGGGGAGAGCGGTTTGGGTATTGGCGCTCTTCCGCTTCCGCTCCTCGCTCACTGACT 5280  
 DB |||||  
 QY 5281 CCGTGGCTCGGTTCGCTGCGGCGAGCGGTATCAGCTCACTCAAAAGCGGTAATAC 5340  
 DB |||||  
 QY 5281 CCGTGGCTCGGTTCGCTGCGGCGAGCGGTATCAGCTCACTCAAAAGCGGTAATAC 5340  
 DB |||||  
 QY 5341 GGTATCCACAGAACTCAGGGGATACGCAAGGAAAGAAACATGTGAGCAAAAGGCCAGCAAA 5400  
 DB |||||  
 QY 5341 GGTATCCACAGAACTCAGGGGATACGCAAGGAAAGAAACATGTGAGCAAAAGGCCAGCAAA 5400  
 DB |||||  
 QY 5401 AGGCCAGGAACCGTAAAAAGCGCGTTCGCTGGCGTTTTTCCATAGGCTCCGCCCCCTG 5460  
 DB |||||  
 QY 5401 AGGCCAGGAACCGTAAAAAGCGCGTTCGCTGGCGTTTTTCCATAGGCTCCGCCCCCTG 5460  
 DB |||||  
 QY 5461 ACCGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAA 5520  
 DB |||||  
 QY 5461 ACCGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAA 5520  
 DB |||||  
 QY 5521 GATACAGCGGTTTCCCGCTGGAAGCTCCCTCGTGGGCTCTCTGTTCCGACCTCGCGC 5580  
 DB |||||  
 QY 5521 GATACAGCGGTTTCCCGCTGGAAGCTCCCTCGTGGGCTCTCTGTTCCGACCTCGCGC 5580  
 DB |||||  
 QY 5581 TTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGCGCTTCTCATAGCTCAC 5640  
 DB |||||  
 QY 5581 TTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGCGCTTCTCATAGCTCAC 5640  
 DB |||||  
 QY 5641 GCTGTAGGTATCTCAGTTCGCTGTAGTGTGCTCGCTCCAAAGCTGGGCTGTGTGCAAGAAC 5700  
 DB |||||  
 QY 5641 GCTGTAGGTATCTCAGTTCGCTGTAGTGTGCTCGCTCCAAAGCTGGGCTGTGTGCAAGAAC 5700  
 DB |||||  
 QY 5701 CCGCGGTTGAGCGGACCGCTGGGCTTATCGGTAACTATCGTCTTGAGTCCAAACCGG 5760  
 DB |||||  
 QY 5701 CCGCGGTTGAGCGGACCGCTGGGCTTATCGGTAACTATCGTCTTGAGTCCAAACCGG 5760  
 DB |||||  
 QY 5761 TAAGACAGGCTTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAGT 5820  
 DB |||||  
 QY 5761 TAAGACAGGCTTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAGT 5820  
 DB |||||  
 QY 5821 ATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGCTTAACCTACGCTACACTAGAGGA 5880  
 DB |||||  
 QY 5821 ATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGCTTAACCTACGCTACACTAGAGGA 5880  
 DB |||||  
 QY 5881 CAGTATTTGGTATCTGCGCTCTGTGAAGCCAGTTACCTTCGGAATAAGAGTTGGTAGCT 5940  
 DB |||||  
 QY 5881 CAGTATTTGGTATCTGCGCTCTGTGAAGCCAGTTACCTTCGGAATAAGAGTTGGTAGCT 5940  
 DB |||||  
 QY 5941 CTTGATCCGGCAAAACAAACCGCTGGTAGCGGTGTTTTTTGTTGCAAGCAGCAGA 6000  
 DB |||||  
 QY 5941 CTTGATCCGGCAAAACAAACCGCTGGTAGCGGTGTTTTTTGTTGCAAGCAGCAGA 6000  
 DB |||||  
 QY 6001 TTACGGCGAGAAAAAGGATCTCAAGAAAGATCCTTTGATCTTTTCTACGGGCTGTGAG 6060  
 DB |||||  
 QY 6001 TTACGGCGAGAAAAAGGATCTCAAGAAAGATCCTTTGATCTTTTCTACGGGCTGTGAG 6060  
 DB |||||  
 QY 6061 CTCAGTGGAAACGAAACTCAGTTTAAAGGATTTGGTCTATGAGATTTATCAAAAAGATCT 6120  
 DB |||||  
 QY 6061 CTCAGTGGAAACGAAACTCAGTTTAAAGGATTTGGTCTATGAGATTTATCAAAAAGATCT 6120  
 DB |||||

QY 6121 TCACCTAGATCCCTTTTAAATTTAAATGAAGTTTTTAAATCAATCTTAAAGTATATATGAGT 6180  
 DB |||||  
 QY 6121 TCACCTAGATCCCTTTTAAATTTAAATGAAGTTTTTAAATCAATCTTAAAGTATATATGAGT 6180  
 DB |||||  
 QY 6181 AAACCTGGTCTGACAGATTACCAATGCTTAAATCAGTGAGGACCACTATCTCAGGAGTCTGTC 6240  
 DB |||||  
 QY 6181 AAACCTGGTCTGACAGATTACCAATGCTTAAATCAGTGAGGACCACTATCTCAGGAGTCTGTC 6240  
 DB |||||  
 QY 6241 TATTTTCGTTTCAATCCATAGTTCGCTGACTCCCGCTCGTGTAGATACTACGATACGGGAGG 6300  
 DB |||||  
 QY 6241 TATTTTCGTTTCAATCCATAGTTCGCTGACTCCCGCTCGTGTAGATACTACGATACGGGAGG 6300  
 DB |||||  
 QY 6301 GCTTACCATCTGGGCCCGAGTGTGCTCAATGATATACCGCGAGACCAAGCTCACCGGCTCCAG 6360  
 DB |||||  
 QY 6301 GCTTACCATCTGGGCCCGAGTGTGCTCAATGATATACCGCGAGACCAAGCTCACCGGCTCCAG 6360  
 DB |||||  
 QY 6361 ATTTATACGCAATAAACCAGCGAGCGGCGAGGCGGAGAGAGTGGTCTCTGCAACTT 6420  
 DB |||||  
 QY 6361 ATTTATACGCAATAAACCAGCGAGCGGCGGAGGCGGAGAGTGGTCTCTGCAACTT 6420  
 DB |||||  
 QY 6421 TATCCGCTCCATCCAGTCTATTAAATGTTGCCGGAAGCTAGAGTAAGTAGTTCGCCAG 6480  
 DB |||||  
 QY 6421 TATCCGCTCCATCCAGTCTATTAAATGTTGCCGGAAGCTAGAGTAAGTAGTTCGCCAG 6480  
 DB |||||  
 QY 6481 TTAATAGTTTGGCAACGTTGTTGCTTACAGGCTCGTGGTGTCAACGCTCGCTGCTT 6540  
 DB |||||  
 QY 6481 TTAATAGTTTGGCAACGTTGTTGCTTACAGGCTCGTGGTGTCAACGCTCGCTGCTT 6540  
 DB |||||  
 QY 6541 TGGTATGCTTCAATTCAGTCCGCTCCGTTCCAAAGTCAAGGGGAGTTACATGATCCCCAT 6600  
 DB |||||  
 QY 6541 TGGTATGCTTCAATTCAGTCCGCTCCGTTCCAAAGTCAAGGGGAGTTACATGATCCCCAT 6600  
 DB |||||  
 QY 6601 GTTGTGCAAAAAGCGGTTAGCTCCTCGTCCGATCGTGTGTCAGAAAGTAAGTTGCG 6660  
 DB |||||  
 QY 6601 GTTGTGCAAAAAGCGGTTAGCTCCTCGTCCGATCGTGTGTCAGAAAGTAAGTTGCG 6660  
 DB |||||  
 QY 6661 GCGAGTGTATCAGCTCATGCTGTTATGCGAGCACTGCATAATCTCTTACTGTGATGCCATC 6720  
 DB |||||  
 QY 6661 GCGAGTGTATCAGCTCATGCTGTTATGCGAGCACTGCATAATCTCTTACTGTGATGCCATC 6720  
 DB |||||  
 QY 6721 GGTAAAGTGTCTTGTGACTGTGAGTACTCAACCAAGTCATTTCTGAGAAATAGTGTAT 6780  
 DB |||||  
 QY 6721 GGTAAAGTGTCTTGTGACTGTGAGTACTCAACCAAGTCATTTCTGAGAAATAGTGTAT 6780  
 DB |||||  
 QY 6781 GCGGCGAGCGAGTTGCTCTTGGCCCGGCTCAATAGCGGATAATACCGGCGCAATAGCAG 6840  
 DB |||||  
 QY 6781 GCGGCGAGCGAGTTGCTCTTGGCCCGGCTCAATAGCGGATAATACCGGCGCAATAGCAG 6840  
 DB |||||  
 QY 6841 AACTTTAAAGTGTCTCATTTGGAAAAACGTTTCTCGGGCGGAAAACTCTCAAGGATCTT 6900  
 DB |||||  
 QY 6841 AACTTTAAAGTGTCTCATTTGGAAAAACGTTTCTCGGGCGGAAAACTCTCAAGGATCTT 6900  
 DB |||||  
 QY 6901 ACCGCTGTGAGATCCAGTTCGATTAACCCACTCGTGCAACCCAACTGATCTTTCAGCATC 6960  
 DB |||||  
 QY 6901 ACCGCTGTGAGATCCAGTTCGATTAACCCACTCGTGCAACCCAACTGATCTTTCAGCATC 6960  
 DB |||||  
 QY 6961 TTTTACTTTTCCAGCGGTTCTGAGTGAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 7020  
 DB |||||  
 QY 6961 TTTTACTTTTCCAGCGGTTCTGAGTGAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAG 7020  
 DB |||||  
 QY 7021 GGGATAAGGGCGACACGGAATGTTGAATCTCATACTCTTCTCTTTTCAATATATTG 7080  
 DB |||||  
 QY 7021 GGGATAAGGGCGACACGGAATGTTGAATCTCATACTCTTCTCTTTTCAATATATTG 7080  
 DB |||||  
 QY 7081 AAGCAATTTATCAGGTTTATGTTCTCATGAGCGGATACATATTGTAATGATTTTAGAAAAA 7140  
 DB |||||  
 QY 7081 AAGCAATTTATCAGGTTTATGTTCTCATGAGCGGATACATATTGTAATGATTTTAGAAAAA 7140  
 DB |||||  
 QY 7141 TAAACAAATAGGGTTCCCGCACATTTCCCGGAAAGTGCCACCTGAGCTCTAGAAAC 7200  
 DB |||||  
 QY 7141 TAAACAAATAGGGTTCCCGCACATTTCCCGGAAAGTGCCACCTGAGCTCTAGAAAC 7200  
 DB |||||  
 QY 7201 CATTATTATCATGACATTAACCTATATAAAATAGGCGTATCACGAGGCGCCCTTTCTGCTCGC 7260  
 DB |||||



Db 7201 CATTATATCATGACATTAACCTATATAAATAGGGGTATACAGAGCCCTTCTGCTCGC 7260  
Qy GCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGACAGTCTCCGGAGACGGTCAACAGC 7320  
Db GCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGACAGTCTCCGGAGACGGTCAACAGC 7320  
Qy TTGTCGTGAAGCGGATGCGGGAGCAGACAGCCGTCAGGCGCGGTGAGGCGGTGTTGG 7380  
Db TTGTCGTGAAGCGGATGCGGGAGCAGACAGCCGTCAGGCGCGGTGAGGCGGTGTTGG 7380  
Qy CCGGTGTCGGGGCTGGCTTAACCTATGCGGCATCAGAGCAGATTTGACTGAGAGTGCACCA 7440  
Db CCGGTGTCGGGGCTGGCTTAACCTATGCGGCATCAGAGCAGATTTGACTGAGAGTGCACCA 7440  
Qy TATCGGTGTGAATAACCGCAGACAGTGTGAAGGAGAAAAATACCGCATCAGGCGCCATTC 7500  
Db TATCGGTGTGAATAACCGCAGACAGTGTGAAGGAGAAAAATACCGCATCAGGCGCCATTC 7500  
Qy GCCATTGAGGCTGGCAACTGTGGGAAGGGGATCGGTGCGGGCTCTTCGCTATTACG 7560  
Db GCCATTGAGGCTGGCAACTGTGGGAAGGGGATCGGTGCGGGCTCTTCGCTATTACG 7560  
Qy CCAGCTGCGGAAAGGGGATGTGTGCAAGCGATTAAAGTTGGGTAAACGCCAGGTTTC 7620  
Db CCAGCTGCGGAAAGGGGATGTGTGCAAGCGATTAAAGTTGGGTAAACGCCAGGTTTC 7620  
Qy CCAGTCACGAGCTGTATAAACGAGCGCCAGTGCC 7654  
Db CCAGTCACGAGCTGTATAAACGAGCGCCAGTGCC 7654

## RESULT 3

US-10-913-288-1

; Sequence 1, Application US/10913288

; Publication No. US20050075492A1

; GENERAL INFORMATION:

; APPLICANT: CHEN, CHANG-ZHENG

; APPLICANT: BARTEL, DAVID

; TITLE OF INVENTION: METHODS AND PRODUCTS FOR EXPRESSION OF microRNA's

; FILE REFERENCE: W00571.70009 US

; CURRENT APPLICATION NUMBER: US/10/913,288

; CURRENT FILING DATE: 2004-08-06

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 1

; LENGTH: 6963

; TYPE: DNA

; ORGANISM: Retroviral

US-10-913-288-1

Query Match 51.0%; Score 3900; DB 19; Length 6963;  
Best Local Similarity 75.8%; Pred. No. 0;  
Matches 5394; Conservative 0; Mismatches 1130; Indels 596; Gaps 22;  
Qy 552 TGAAGACCCACCTGTAGTTGGCAAGCTAGCTTAAGTAACGCCATTTTGCAGGCAAT 611  
Db 1 TGAAGACCCACCTGTAGTTGGCAAGCTAGCTTAAGTAACGCCATTTTGCAGGCAAT 60  
Qy 612 GGAATAACATACTGAGAAATAGAGAGTTTCAAGTCAAGTTAGGAACAGAGACAGCA 671  
Db 61 GGAATAACATACTGAGAAATAGAGAGTTTCAAGTCAAGTTAGGAACAGAGACAGCA 119  
Qy 672 GAATATGGGCCAAACAGGATATCTGTGTAAGCAGTTTCTCCCGCTCAGGCGCAAGAA 731  
Db 120 ----- 119  
Qy 732 CAGTTGGAAACAGGAAATATGGCCAAACAGGATATCTGTGTAAGCAGTTTCTCCCGCG 791  
Db 120 -----AGAAATGGGCCAAACAGGATATCTGTGTAAGCAGTTTCTCCCGCG 166  
Qy 792 GCTCAGGCCCAAGAACAGATGGTCCCGAGATGCGGTCCCGCCCTCAGCAGTTTCTAGAGA 851

Db 167 GCTCAGGCCCAAGAACAGATGGTCCCGAGATGCGGTCCCGCCCTCAGCAGTTTCTAGAGA 226  
Qy ACCATCAAGATGTTTCCAGGGTGCCCAAGAGACCTG- AAATGACCTCTGTGCTTATTGAA 910  
Db ACCATCAAGATGTTTCCAGGGTGCCCAAGAGACCTGAAAATGACCTGTGCTTATTGAA 286  
Qy CTAAACCAATCAGTTTCTGCTTCTGTTTCGGCGCTTCTGCTCCCGAGCTCAATAA 970  
Db CTAAACCAATCAGTTTCTGCTTCTGTTTCGGCGCTTCTGCTCCCGAGCTCAATAA 346  
Qy AAGAGCCCAACACCCCTCACTCGGCGGCCAGTCTCTCCGATAGACTGCGTCCCGCGGTA 1030  
Db AAGAGCCCAACACCCCTCACTCGGCGGCCAGTCTCTCCGATAGACTGCGTCCCGCGGTA 406  
Qy CCGTATTTCCCAATAAAGCCTCTTGTGTCATTCGAAATCGTGGAGCTCGCTGATCCTT 1090  
Db CCGTATTTCCCAATAAAGCCTCTTGTGTCATTCGAAATCGTGGAGCTCGCTGATCCTT 466  
Qy GGGAGGGTCTCTCAGATTGATTCAGTCCCACTCGGGGTCTTTCATTTGGAGGTTCC 1150  
Db GGGAGGGTCTCTCAGATTGATTCAGTCCCACTCGGGGTCTTTCATTTGGAGGTTCC 526  
Qy ACCGAGATTTGGAGACCCCTGCCCCAGGACCAACCGACCCCGCGGAGGTAAAGCTGG 1210  
Db ACCGAGATTTGGAGACCCCTGCCCCAGGACCAACCGACCCCGCGGAGGTAAAGCTGG 586  
Qy CCAGCAACTATCTGTCTGTCTGTC-----CGATTGTCTAGTGTCTATGACTGATTTTA 1262  
Db CCAGCGTCTGTTCTGTCTGTCTGTCCTGTTGTCGCTGTTGTCGCGCATCTAATGTT 646  
Qy TGCGCCTCGCTCGGTACTAGTACTAGTCTCTGATCTGCTGCGGACCCGCTGCTGAA 1322  
Db TGCGCCTCGCTCTGACTAGTACTAGTCTCTGATCTGCTGCGGACCCGCTGCTGAA 706  
Qy CTGACGAGTTTCGAAACACCCCGCGCAACCTCGGAGACGCTCCAGGAGCTTCGGGGGCC 1382  
Db CTGACGAGTTTCGAAACACCCCGCGCAACCTCGGAGACGCTCCAGGAGCTTCGGGGGCC 766  
Qy GTTTTGTGCGCCGACCTGAGTCTCTAAATCCCGATCGTTTGGAGCTCTTTGGTGCACCC 1442  
Db GTTTTGTGCGCCGACCTGAGGAGGAGTCTGATGTGGAATCCGACCC----- 814  
Qy CCCTTAGAGAGGGATATGTTCTGCTAGGAGACGAGAACCTTAAACAGTTCCCGCT 1502  
Db -----CGTCAGGATATGTTCTGCTAGGAGACGAGAACCTTAAACAGTTCCCGCT 867  
Qy CCGTCTGAAATTTTGTGTTTGGACCGAAGCGCGCGCGCTTGT---CTGC 1559  
Db CCGTCTGAAATTTTGTGTTTGGAAACCGAAGCGCGCGCTTGTCTCTGCTGCGCGC 927  
Qy TGCAAGCATCGTTCTGTTGTTCTCTGCTGATGTTTCTGTTTCTGTTTCTGTTTCTGTT 1619  
Db TGCAAGCATCGTTCTGTTGTTCTCTGCTGATGTTTCTGTTTCTGTTTCTGTTTCTGTT 987  
Qy GGCCGGGCTAGACTGTTTCACTCCCTTAAGTTTGAACCTTAGTCTCACTGAGAAAGATGTC 1679  
Db GGCC-----AGACTGTTTACCCTTAAAGTTTGAACCTTAGTCTCACTGAGAAAGATGTC 1041  
Qy GAGCGGATCGCTCAACACAGTCTGAGTCTCAAGAGAGACGCTTGGTTTACCTTCTGC 1739  
Db GAGCGGATCGCTCAACACAGTCTGAGTCTCAAGAGAGACGCTTGGTTTACCTTCTGC 1101  
Qy TCTCAGAAATGGCCAAACCTTTAAGCTCGGATGGCGCGGAGACGCGACCTTTAACCGAGAC 1799  
Db TCTCAGAAATGGCCAAACCTTTAAGCTCGGATGGCGCGGAGACGCGACCTTTAACCGAGAC 1161  
Qy CTATATCAACAGGTTAAGTCAAGGCTCTTTTCACTGCGCGCATGAGACCCAGACAG 1859  
Db CTATATCAACAGGTTAAGTCAAGGCTCTTTTCACTGCGCGCATGAGACCCAGACAG 1221  
Qy GTCCCTTACATCTGAGCTCGGAGACCTTGTGCTTTTGAACCCCTCTCCCTGGTCAAGCCC 1919  
Db GTCCCTTACATCTGAGCTCGGAGACCTTGTGCTTTTGAACCCCTCTCCCTGGTCAAGCCC 1281







QY 6296 GGAGGGCTTACCATCTGGCCCGAGTGTCTCAATGATACCGCGAGACCCACGCTCACCGGC 6355  
 DB 5182 GGAGGGCTTACCATCTGGCCCGAGTGTCTCAATGATACCGCGAGACCCACGCTCACCGGC 5241  
 QY 6356 TCCAGATTTATCAGCAATAAACCAGCAGCCGGAAGGCGGAGCGCAGAGTGGTCTCTGC 6415  
 DB 5242 TCCAGATTTATCAGCAATAAACCAGCAGCCGGAAGGCGGAGCGCAGAGTGGTCTCTGC 5301  
 QY 6416 AACTTTATCCGCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTC 6475  
 DB 5302 AACTTTATCCGCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTC 5361  
 QY 6476 GCCAGTTAATAGTTTGGCAACGTTGTTGCCAATGCTACAGGC-TGCTGGTGTCAAGCTC 6534  
 DB 5362 GCCAGTTAATAGTTTGGCAACGTTGTTGCCAATGCTACAGGC-TGCTGGTGTCAAGCTC 5421  
 QY 6535 GTCTGTTGATAGTCTTCAATCAGCTCCGGTTCCTCAACGATCAAGCGAGTTACATGATC 6594  
 DB 5422 GTCTGTTGATAGTCTTCAATCAGCTCCGGTTCCTCAACGATCAAGCGAGTTACATGATC 5481  
 QY 6595 CCCCATGTTGTGCAAAAAAGCGTTAGCTCTCTCGGTCTCTCGATGCTGTTGTGCAAGTAA 6654  
 DB 5482 CCCCATGTTGTGCAAAAAAGCGTTAGCTCTCTCGGTCTCTCGATGCTGTTGTGCAAGTAA 5541  
 QY 6655 GTTGGCGCAGTGTATCACTCATGTTATGTCAGCACTGCATATTTCTTACTGTCTAT 6714  
 DB 5542 GTTGGCGCAGTGTATCACTCATGTTATGTCAGCACTGCATATTTCTTACTGTCTAT 5601  
 QY 6715 GCCATCCGTAAGATGCTTTCTGTGACTGCTGAGTACTCAACCAAGTCAATCTCGAATA 6774  
 DB 5602 GCCATCCGTAAGATGCTTTCTGTGACTGCTGAGTACTCAACCAAGTCAATCTCGAATA 5661  
 QY 6775 GTGTATGCGCGACCGAGTTGCTCTTGGCGGTCATATCGGGAATAATACCGCGCCACA 6834  
 DB 5662 GTGTATGCGCGACCGAGTTGCTCTTGGCGGTCATATCGGGAATAATACCGCGCCACA 5721  
 QY 6835 TAGCAGAACTTTAAAGTCTCATCATTTGGAAAGCTTCTTGGGGCGAAACTCTCAG 6894  
 DB 5722 TAGCAGAACTTTAAAGTCTCATCATTTGGAAAGCTTCTTGGGGCGAAACTCTCAG 5781  
 QY 6895 GATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGACCCCACTGATCTTC 6954  
 DB 5782 GATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGACCCCACTGATCTTC 5841  
 QY 6955 AGCATCTTTTATTTTACCGAGTCTTCTGGGTGAGCAAAAAAGGAGGCAAAATGCCGC 7014  
 DB 5842 AGCATCTTTTATTTTACCGAGTCTTCTGGGTGAGCAAAAAAGGAGGCAAAATGCCGC 5901  
 QY 7015 AAAAAAGGAATAAGGGGACACGGAATGTTGAATCTCATCTCTCTTTTCAATA 7074  
 DB 5902 AAAAAAGGAATAAGGGGACACGGAATGTTGAATCTCATCTCTCTTTTCAATA 5961  
 QY 7075 TTATTGAAGCATTTATCAGGTTATTGTTCTCATGAGCGATACATATTTGAATGATTTA 7134  
 DB 5962 TTATTGAAGCATTTATCAGGTTATTGTTCTCATGAGCGATACATATTTGAATGATTTA 6021  
 QY 7135 GAAAAATAAACAAATAGGGTTCCGCGCACATTTCCCGGAAAGTGCCACCTGACGTCTA 7194  
 DB 6022 GAAAAATAAACAAATAGGGTTCCGCGCACATTTCCCGGAAAGTGCCACCTGACGTCTA 6081  
 QY 7195 AGAAACCATTTATATCAATGACATTAACCTATAAAAAATAGCGTATCACGAGCCCTTCG 7254  
 DB 6082 AGAAACCATTTATATCAATGACATTAACCTATAAAAAATAGCGTATCACGAGGCCCTTCG 6141  
 QY 7255 TCTCGCGCTTTTCGGTGTAGCAGGTGAAACCTCTTGACACATGACGCTCCCGGAGCGT 7314  
 DB 6142 TCTCGCGCTTTTCGGTGTAGCAGGTGAAACCTCTTGACACATGACGCTCCCGGAGCGT 6201  
 QY 7315 CACAGCTTCTGTAAAGCGGATGCCGGGAGCAGCAAGCCCGTCAAGGCGCGTCAAGCGGG 7374  
 DB 6202 CACAGCTTCTGTAAAGCGGATGCCGGGAGCAGCAAGCCCGTCAAGGCGCGTCAAGCGGG 6261

QY 7375 TGTGGCGGTGTCTGGGCTGTCTAACTATCGCGCATCAGACGAGATTGTACTGAGACT 7434  
 DB 6262 TGTGGCGGTGTCTGGGCTGTCTAACTATCGCGCATCAGACGAGATTGTACTGAGACT 6321  
 QY 7435 GCACATATGCGGTGTGAAATACCGCACAGATGCGTAAAGGAGAAATACCGCATCAGCG 7494  
 DB 6322 GCACATATGCGGTGTGAAATACCGCACAGATGCGTAAAGGAGAAATACCGCATCAGCG 6381  
 QY 7495 CCAATTGCGCATTCAGGCTGCGCAACTGTTGGGAAGGCGCATCGGTGCGGGCTCTTCGCT 7554  
 DB 6382 CCAATTGCGCATTCAGGCTGCGCAACTGTTGGGAAGGCGCATCGGTGCGGGCTCTTCGCT 6441  
 QY 7555 ATTACGCGAGCTGGCGGAAAGGGGATGCTCTCAAGGCGCATTAAGTTGGGTAAACGCGCAG 7614  
 DB 6442 ATTAAGCGAGCTGGCGGAAAGGGGATGCTCTCAAGGCGCATTAAGTTGGGTAAACGCGCAG 6501  
 QY 7615 GTTTTCCAGTCCACAGCTGTTTAAACGACGCGCCAGTGCC 7654  
 DB 6502 GTTTTCCAGTCCACAGCTGTTTAAACGACGCGCCAGTGCC 6541

RESULT 4  
 US-10-789-938B-3  
 ; Sequence 3, Application US/10789938B  
 ; Publication No. US20050009180A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Yang, Lili  
 ; APPLICANT: Van Parijs, Luk  
 ; APPLICANT: Baltimore, David  
 ; TITLE OF INVENTION: METHOD FOR THE GENERATION OF  
 ; TITLE OF INVENTION: ANTIGEN-SPECIFIC LYMPHOCYTES  
 ; FILE REFERENCE: CALTE.008CP1  
 ; CURRENT APPLICATION NUMBER: US/10/789,938B  
 ; PRIOR FILING DATE: 2004-02-27  
 ; PRIOR APPLICATION NUMBER: 10/317,078  
 ; PRIOR FILING DATE: 2002-12-10  
 ; PRIOR APPLICATION NUMBER: 60/394,803  
 ; PRIOR FILING DATE: 2002-07-08  
 ; PRIOR APPLICATION NUMBER: 60/339,375  
 ; PRIOR FILING DATE: 2001-12-10  
 ; NUMBER OF SEQ ID NOS: 3  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 3  
 ; LENGTH: 7277  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: This represents a retroviral vector encoding a  
 ; OTHER INFORMATION: T-cell receptor that recognizes an epitope of  
 ; OTHER INFORMATION: Mart-1.  
 US-10-789-938B-3

Query Match 43.9%; Score 3359.8; DB 19; Length 7277;  
 Best Local Similarity 69.4%; Pred. No. 0;  
 Matches 4952; Conservative 0; Mismatches 2022; Indels 159; Gaps 22;

QY 552 TGAAGACCCACCTGTAGTTGGCAAGCTAGTCTAGTAACGCCATTTTGAAGGCAT 611  
 DB 1 TGAAGACCCACCTGTAGTTGGCAAGCTAGTCTAGTAACGCCATTTTGAAGGCAT 60  
 QY 612 GGAATAATACATAACCTAGTAAGAGTTTCAAGTCAAGAGTTAGGAACAGAGAGACGCA 671  
 DB 61 GGAATAATACATAACCTAGTAAGAGTTTCAAGTCAAGAGTTAGGAACAGAGAGACGCA 119  
 QY 672 GAATATGGGCCAAACAGGATATCTGTGTGAAGCAGTTCTCTGCCCGCTCAGGCCCAAGAA 731  
 DB 120 ----- 119  
 QY 732 CAGTTGGAAACAGGAGAAATATGGGCCAAACAGGATATCTGTGTGAAGCAGTTCTCTGCCCG 791  
 DB 120 -----AGAATATGGGCCAAACAGGATATCTGTGTGAAGCAGTTCTCTGCCCG 166  
 QY 792 GCTCAGGCCCAAGAACAGATGTTGTCGCCGCTCCCGCTCAGCAGTTTCTTAGAGA 851

Db 167 GTTGGGCGAAGACAGATGTTCCAGATCGGTCCCGCCTCAGACGTTCTAGAGA 226  
 QY 852 ACCATCAGATGTTCCAGGGTCCCAAGGACCTG-AAATGACCCCTGTCCTTATTTGAA 910  
 Db 227 ACCATCAGATGTTCCAGGGTCCCAAGGACCTGAAATGACCTGTCGCTTATTTGAA 286  
 QY 911 CTAACCAATCAGTTGCTGCTCTGCTTCTGTTGCGGCTTCTGCTCCCGAGCTCAATAA 970  
 Db 287 CTAACCAATCAGTTGCTGCTCTGCTTCTGTTGCGGCTTCTGCTCCCGAGCTCAATAA 346  
 QY 971 AAGAGCCACAAACCCCTCACTCGGCGCGCAGTCTCCGATAGACTGGTGGCCGGGTA 1030  
 Db 347 AAGAGCCACAAACCCCTCACTCGGCGCGCAGTCTCCGATAGACTGGTGGCCGGGTA 406  
 QY 1031 CCGGTATTTCCCAATAAAGCCTCTTGTGTTTGGCATCCGAATCGTGGACTCGCTGATCCTT 1090  
 Db 407 CCGGTATTTCCCAATAAAGCCTCTTGTGTTTGGCATCCGAATCGTGGACTCGCTGATCCTT 466  
 QY 1091 GGGAGGTCCTCAGATGATTGACTGCCACCTCGGGGTCTTTCAATTTGGAGGTTCC 1150  
 Db 467 GGGAGGTCCTCAGATGATTGACTGCCACCTCGGGGTCTTTCAATTTGGAGGTTCC 526  
 QY 1151 ACCGAGATTTGGAGACCCCTGCGCAGGACCAACCGACCCCGCGGGAGGTAAGCTGG 1210  
 Db 527 ACCGAGATTTGGAGACCCCTGCGCAGGACCAACCGACCCCGCGGGAGGTAAGCTGG 586  
 QY 1211 CCAGCAACTTATCTGTGTCGTG-----CGATTGCTAGTGTCTATGACTGATTTTA 1262  
 Db 587 CCAGCGTCGTTTCTGTGTCGTCTGCTGTTGTCGTTGTCGCGCATCAATGTT 646  
 QY 1263 TGGCCTCGCTGGTACTAGTGTAACTAGTCTGTATCTGCGGACCCGCTGGTGAA 1322  
 Db 647 TGGCCTCGCTGTACTAGTGTAACTAGTCTGTATCTGCGGACCCGCTGGTGAA 706  
 QY 1323 CTGAGGTCGGAACACCGGCGCAACCTGGGAGACGTCGAGGACTTCGGGGCC 1382  
 Db 707 CTGAGGTCGGAACACCGGCGCAACCTGGGAGACGTCGAGGACTTCGGGGCC 766  
 QY 1383 GTTTTGTGGCCGACCTGAGTCTTAAATCCGATCCGTTTGGTGTGACCC 1442  
 Db 767 GTTTTGTGGCCGACCTGAGGAGGAGTGTGATGTGAAATCCGACCC----- 814  
 QY 1443 CCTTAGAGGAGGATATGTTGTTCTGTGAGACGAGAACTTAAACAGTTCGCGCT 1502  
 Db 815 -----CGTCAGGATATGTTGTTCTGTGAGACGAGAACTTAAACAGTTCGCGCT 867  
 QY 1503 CCGTCTGAATTTTGTCTTGGTTTGGACCGAGCCGCGCGCTTGTGTC---TGC 1559  
 Db 868 CCGTCTGAATTTTGTCTTGGTTTGGAAACGAAAGCCGCGCTTGTCTGCGAGCGC 927  
 QY 1560 TGCAGCATCGTTCTGTGTTGTTCTGTCTGACTGTGTTTCTGTATTTCTGAAATAATG 1619  
 Db 928 TGCAGCATCGTTCTGTGTTGTTCTGTCTGACTGTGTTTCTGTATTTGTCTGAAATAATG 987  
 QY 1620 GGCCGGGCTAGACTGTTTACCTCCCTTAAAGTTTGAACCTTAGTCTAGTCACTGAAAGATGTC 1679  
 Db 988 GGCC-----AGACTGTTTACCCTTCCCTTAAAGTTTGAACCTTAGTCTAGTCACTGAAAGATGTC 1041  
 QY 1680 GAGCGATCGCTCACAACAGTCCGTAGATGTCAAGAAGACGTTGGTTTACCTTCTGC 1739  
 Db 1042 GAGCGATCGCTCACAACAGTCCGTAGATGTCAAGAAGACGTTGGTTTACCTTCTGC 1101  
 QY 1740 TCTGAGAAATGGCAACCTTTAAAGTCCGATGGCCGAGACGCGACCTTTAAACCGAGAC 1799  
 Db 1102 TCTGAGAAATGGCAACCTTTAAAGTCCGATGGCCGAGACGCGACCTTTAAACCGAGAC 1161  
 QY 1800 CTCATCACCCAGGTTAAGATCAAGTCTTTTCACTGGCCGCGATGACACCCAGACGAG 1859  
 Db 1162 CTCATCACCCAGGTTAAGATCAAGTCTTTTCACTGGCCGCGATGACACCCAGACGAG 1221  
 QY 1860 GTCCCTTACATCGTGACCTCGGAAGCTTGGGTTTGAACCCCTTCCCTGGGTCAAGGCC 1919

Db 1222 GTCCCTACATCGTGACCTGGGAAGCCTTGGCTTTTGTAGACCCCTCCCTCGGTCAGGCC 1281  
 QY 1920 TTTGTACACCTTAAGCCTCCGCTCCTCTTCTCCTCATCGCCCGCTCTCTCCCTTTGAA 1979  
 Db 1282 TTTGTACACCTTAAGCCTCCGCTCCTCTTCTCCTCATCGCCCGCTCTCTCCCTTTGAA 1341  
 QY 1980 CTTCTCTGTTTGAACCCCGCTCGATCTCTCTTATCAGCCCTCACTCTCTCTAGGC 2039  
 Db 1342 CTTCTCTGTTTGAACCCCGCTCGATCTCTCTTATCAGCCCTCACTCTCTCTAGGC 1401  
 QY 2040 GCCCCATATGGCCATATGAGATCTTATATGGGCAACCCCGCTCTTGTAACTTCCCT 2099  
 Db 1402 GCGGAGATCTCTGAGGT-----TAAAGAAAGGATCAATGTTCTTGAACATTTA 1455  
 QY 2100 GACCTGATGACAAAGAGTTACTAAAGCCCTCTCTCCAAGCTCACTTACAGGCTTCT 2159  
 Db 1456 ATAATCTTGTGATGTCAGCTGACGTGAGTGGTCAAGAGTGAATCAGAGTCTCTCA 1515  
 QY 2160 ACTTAGTCCAGACGAACTCTGGAGACCTCTGGCGGAGCCCTACCAAGAACACTGGAGC 2219  
 Db 1516 TCTATGTTTATCCAGGAAGGAGAGATGTCTCCATGAATCTGCACTTCTTCAAGCATATTT 1575  
 QY 2220 GACCGTGG-TACCTCACCTTTACCGAGTCCGCGACACAGTGGGTCCGCGACCCAG 2278  
 Db 1576 AACCTGGCTATGGTACAAGAGGACCTCGGGAAGGTCTCTCTTGTATAGCTTTA 1635  
 QY 2279 ACTAAGAACCTTAGAACCTCGCTGGAAAGGACCTTACACAGTCTCTGACACCCGCCAC 2338  
 Db 1636 TATAAGGCTGTGAATTGACCTCAAAATGGAAGACTGACTGCTCAGTTTGGTATTAACCA 1695  
 QY 2339 GCCCTCAAGTAGAGCGGATCGAGCTTGGATACAGCGCCGCCAGTGAAGGCTCCGAC 2398  
 Db 1696 AAGGACAGCTTCTGAAATATCTCAGCATCCATACCTAGTGTAGTGGCAT-----CTAC 1749  
 QY 2399 CCGGGGTGGACCATCTCTAGACTGCTCATGGGATGGAGTGTATCATCTCTTCTTGG 2458  
 Db 1750 TTCTGTCTGTGGGACCGGTAAACAGTTCTATTTTGGGACAGGACAGTTTGAAGTTC 1809  
 QY 2459 TAGCAACAGCTACAGGTGTCCACCTCCGACATCCAGCTGACCCAGAGCCCAAGCAGCTGA 2518  
 Db 1810 ATTCCAAATATCCAGAACCTTGACCTCCCTGCTGCTGACAGCTGAGAGACTCTAAATCCAGT 1869  
 QY 2519 GCGCCAGCGTGGGTGACAGAGTGACCATCACTGTGAAGGCGAGTCAAGATGTGGTACTT 2578  
 Db 1870 GA-----CAAGTCTGTCTGCTTATCCAGGATTTGATTTCTCAACAATAATGTGTACAAA 1924  
 QY 2579 CTGTAGCTTGTGATCCAGCAGAGGACGATAGGCTCCAAAGCTGCTGATCTACTGGACAT 2638  
 Db 1925 GTAAGGATTTGTATGTATATCAAGACAAAATGTGTCTAGACATGAGGTCTATGGACT 1984  
 QY 2639 CCACCCGCGACACTGGGTGTGCAAGCAGATTTCAAGGTTAGCGGTAGCGGTACCGACTTCA 2698  
 Db 1985 TCAAGAGCAACAGTGTGTGGCTGGAGCAACAAATCTGACTTTGCTATGTGCAACCGCT 2044  
 QY 2699 CTTTCAACATCAGCAGCTCCAGCCAGAGGACATCGCCACCTACTACTGCGCAGCAATATA 2758  
 Db 2045 TCAACAACAGCATTTATCCAGAGACACCTTCTTCCCGAGCCAGAAAGTTCTCTGTGATG 2104  
 QY 2759 GCCTCTATCGTCTGTGGCCCAAGGACCAAGTGGAAATCAACAGAGTGGCTCAGGAT 2818  
 Db 2105 TCAAGCTGGTCGAGAAAGCTTTGAAACAGATACGAACCTTAACTTTTCAAAACCTGTGAG 2164  
 QY 2819 CGGTGGATCCGG--CTCTGTGGCTCAGGATCGAGGTCCAACTGTGGGAGAGCGGTGG 2876  
 Db 2165 TGATTTGGTTCGAAATCCTCTCTGAAAGGTGGCGGGTTTAACTGCTCATGACGCTGC 2224  
 QY 2877 AGGTGTTGTGCAACCTGGCCGCTCTCTGCGCTGTCTGCTCCGATCTGGCTTGGCTT 2936  
 Db 2225 GGCTGTGGTCCAGCTGAGAAATTCGCTGAGGCTGTGGTCCAGTTGACGCCCTCTCCCTC 2284  
 QY 2937 CACCAATATTTGGATGAGTTGGGTGAGAC-----AGGCACCTTGGAAAGAGTCTTGTAGTGA 2992  
 Db 2285 CCCCCCCCCTAACGTTTACTGGCCGAGCGCTTGGAAATAGGCGGTTGGTGTGTTGTCTA 2344











Db 1222 GTCCCTACATCGTGAACCTGGGAAGCCCTTGGCTTTTGACCCCCCTCCCTGGGTCAAGCCC 1281  
Qy 1920 TTTGTACACCTTAAGCCTCCGCTCCTCTCTTCCATCCGCCCCCTCTCTCCCTTGA 1979  
Db 1282 TTTGTACACCTTAAGCCTCCGCTCCTCTCTTCCATCCGCCCCCTCTCTCCCTTGA 1341  
Qy 1980 CCTCCTCTTTGACCCCGCTCGATCTCCTCTTTATCCAGCCCTCACTCTCTCTAGGC 2039  
Db 1342 CCTCCTCTTTGACCCCGCTCGATCTCCTCTTTATCCAGCCCTCACTCTCTCTAGGC 1401  
Qy 2040 GCCCCATAT-----GGCCATATGATCTTATATGGGCGACCCCGCCCTTGTAACT 2094  
Db 1402 GCGGATCTCTCGAGGTTAAGAAACCGATCCATGGTGAAGATCCGGCAATTTTGTG 1461  
Qy 2095 TCCCTGACCCCTGACATGATAAGAGTTACTAACAGCCCTCTCTCCAAAGCTCACTTACAGG 2154  
Db 1462 GCTATTTTGTGCTTCAGCTAAGCTGTGTAACTGCGCCGCAAAATGAAGTGGACAGAT 1521  
Qy 2155 CTT-----CTACTTAGTTCAGACGAAGTCTGGAGACCTCTGGGGGAGCCCTACCAAGAA 2209  
Db 1522 CCTCAGAACCTGACTGCGCCAGGAAGGAATTTATCAATCAACTGACGATTACTCGGTA 1581  
Qy 2210 CAACCTGGACCGCTGCTACCTCACTTACCGAGTCCGGGACACAGTGTGGTCCGC 2269  
Db 1582 GGAATAAGTGCTTACACTGGCTGCAACAGCATCCAGGAGGGAATTTGTTCTCTTGT 1641  
Qy 2270 CGACACACAGACTAAGAACTCGCTGGAAGGACCTTACACAGTCTCTCTGACC 2329  
Db 1642 ATGTGAGCTCAGGGAAGAGCATGGAAGTTAATGGCAATAAATACAGGAA 1701  
Qy 2330 ACCCCACCGCCCTCAAAAGTGAAGGCAATCGAGCTTGGATACAGCGCCGCCAGTGAAG 2389  
Db 1702 AAGCACAGCTCCCTGCACATCACAGCTCCCATCCAGAGACTCTGCGCTCTACATCTGT 1761  
Qy 2390 GCTGCGGACCCGGGGTGGACCATCTCTAGACTGCCATGGATGGAGCTGTATCATCC 2449  
Db 1762 GCTGCCCTCATTAATTCAGGAGC--CCAGAAAGTGGTATTTGGCCAAAGAAACAGGCTGA 1819  
Qy 2450 TCTTCTTGGTAGCAACAGCTACAGGTGTCCTCGACATCCAGCTGACCCAGAGCCCAA 2509  
Db 1820 CTATCAACCCAAATATCCAGAACCTGACCTGCGCTGTACAGCTGAGAGCTTAAT 1879  
Qy 2510 GCAGCTGAGCGCCAGCGTGGGTGAAGAGTGAACCATCACTGTAAGGCCAGTCAGGATG 2569  
Db 1880 CCAG-----TGACAAGTCTGTCTGCCTATTACCCGATTTTGAATCTCAAAACAATG 1930  
Qy 2570 TGGGTACTTCTGTAGCTTGGTACAGAGAGAGCCAGGTAAAGCTCCAAAGCTGTATCT 2629  
Db 1931 TGTCACAAAAGTAAGGATTTCTGATGTGTATATCACAGACAAAACCTGTGTAGACATGAGGT 1990  
Qy 2630 ACTGGACATCCACCCGACACTGTGTGTCGAAGCAGATTACAGCGTAGCGGTAGCGTA 2689  
Db 1991 CTATGGACTTCAGAGCAACAGTCTGTGGCTGTGAGCAAAATCTGACTTTGCATGTG 2050  
Qy 2690 CCGACTTCACCTTCAACATCAGCAGCTCCAGCCAGAGGACATGCCACCTACTACTGCC 2749  
Db 2051 CAACGCTTCAACAAACAGCATTTATCCAGAAGACACCTCTTCCCCAGCCAGAAAGTT 2110  
Qy 2750 AGCAATATAGCCTCTATCGGTGTTCCGGCCAAAGGGAACCAAGGTGAATCAACAGAGTG 2809  
Db 2111 CCTGTGATGTCAAGCTGTGAGAAAGCTTTGAAACAGATACGAAACCTTAACATTTCAA 2170  
Qy 2810 GCTCAGGATCGGTGGATCCGG--CTCTGTGGCTCAGGATCGGAGTCCAACTGGTGA 2867  
Db 2171 ACCTGTAGTGAATGGGTTCGAATCTCTCTCTGAAAGTGGCCGGGTTTAATCTGTCTCA 2230  
Qy 2868 GAGCGGTGGAGGTGTGTGCAACCTGGCCGCTCCTGCGCTGTCTCTCCGATCTGG 2927  
Db 2231 TGACGCTCGGCTGTGGTCCAGCTGAGAAATTCGCTGAGGCTGTGGTCCAGTTGACGCC 2290  
Qy 2928 CTTGATTTTCAACCATATTCGATAGTTGGTGGAGACGACCTGGAAAGGTCTTGA 2987

Db 2291 CTCCTCCCTCCCCCCCCCTAAAGTTACTGCGCGGAAGCCG-----CTTGAATAAGCCCGT 2346  
Qy 2988 GTGGATTGGAGAAATTAATCCAGATAGCAGTACGATTAACATATATGCGCGTCTCTAAAGAA 3047  
Db 2347 GTGCGTTGTCTATATGTTATTTTCCACCAATATGCGCGTCTTTTGGCAATGTG---AGGG 2403  
Qy 3048 TAGATTTACAATATCGGAGACAAGCCCAAGAAACATTTGTTCTCTGCAAAATGGACAGCT 3107  
Db 2404 CCGGAAACCTTGGCCCTGTCTTCTTGACGAGCATTTCTTAGGGGTCTTTCCCTCTCGCA 2463  
Qy 3108 GAGACCCGAGACACCGGGTCTATTTTGTGGAAGCCTTTACTTTCGGCTTCCCTGTGTT 3167  
Db 2464 AAGGAATGCAAGTCTGTGTAATGTGTGAAGGAAGCATGTTCTCTGAAGCTTCTTGA 2523  
Qy 3168 TGCTTATTTGGGGCAAGGGACCCCGGTCAACGTCTCCAGTCTTAAGCCCAACCAAGACCC 3227  
Db 2524 GACAAAACAAGCTGTAGCGACCTTTGCAGS---CAGCGGAACCCCACTTGCAGCA 2580  
Qy 3228 AGCGCGGGACCAACAACCGGCGCCCAATCGGTGCGAGCCCTGTCTCTGCGGCC 3287  
Db 2581 GGTGCTCTTGGCGCAAAAGCCAGCTGTATGAATACACCTGCAAGCGGCAACAACCC 2640  
Qy 3288 AGAGCGGCTCGGCCAGCGCGGGCGCAGTGCACACGAGGGGGCTTTCGCCCT 3347  
Db 2641 AGTCCAGCTTGTGAGTTGGATAGTTGTGGAAGAGTCAAAATGGCTCTCTCAAGCGTAT 2700  
Qy 3348 GGATCCCAAACTCTGCTACCTGCTGGATGGAACTCTCTTCTATCTATGTGTCATCTCAC 3407  
Db 2701 TCAACAAGGGGCTGAAGATGCCAGAGGTACCCCATTTGTATGGATCTGATCTGGGGC 2760  
Qy 3408 TGCCTTGTCTGAGAGTGAAGTTACAGAGGCGCAGAGCCCCCGCTTACCAGCAGGG 3467  
Db 2761 CTCGTGTGCATGCTTTACATGTTAGTCGAGTTTAAAAAAGCTAGGCCCCCCGAA 2820  
Qy 3468 CCAGAACCGCTCTATAACAGCTCAATCTAGGAGGAGAGAGTACGATGTTTGA 3527  
Db 2821 CCAGGGGACGTGTTTCTTTTGAATAACAGATGAATATGAGCCCAACCACTGGA 2880  
Qy 3528 CAAGAGAGCTGGCGGACCTGAGATGGGGGAAAGCCGAGAAAGAAACCCCTCAGA 3587  
Db 2881 CCTGGACTTCTGTG-----TGTCCTTTGATCTCTGTAGGAGCATACAGATGC 2935  
Qy 3588 AGGCTGTACAATGAACCTGCAGAAAGATAAGATGGCGAGCCCTACAGTGAATGGAT 3647  
Db 2936 TGGAGTTATCCAGTCAACCCGCTCAGGTGACAGAGATGGGCAAGAAGTGACT---CT 2992  
Qy 3648 GAAAGGCGAGCGCGGAGGGCAAGGGCAGATGGCCCTTACCAGGCTCTCAGTACAGC 3707  
Db 2993 GAGATGTAAACCAATTTGAGGCCCAAACTCCCTTTTCTGGTACAGACAGACATGATCGG 3052  
Qy 3708 CACCAAGGACACCTACGACGCTTTCATGTCAGGCCCCCTGCCCCCTGCTAACTCGAGC 3767  
Db 3053 GGGACTGGATGTGCTCATTTACTTTTAAACAACAGTTCCGATAGTATTTCA---GGAT 3109  
Qy 3768 GGCGCGGATCCGGATTAAGTCAATTTTGTAAAGACAGGATATCAGTGGTCCAGGCTCTA 3827  
Db 3110 GCCCGAGATCGATCTCTCAGCTAAGTGCCTAATGTCATCTCTCCACTCTGAAGATCCA 3169  
Qy 3828 GTTTTGACTCAAAATATCACCAGCTGAAGCTATAGAGTACGAGCCATAGATAAATAA 3887  
Db 3170 GCCCTCAAGCCAGGAGCTCAGCTGTGTACTTCTGTGCCAGCAGCCCCGGGGCAATGA 3229  
Qy 3888 AAGATTTTATTTAGTCTCCAGAAAAAGGGGGAATGAAAGACCCCACTGTAGGTTTGGC 3947  
Db 3230 GCAGTCTTCTGGGCCAGGACACGGCTCACCGTCTAGAGGACCTGAAAAAGCTGTCTCC 3289  
Qy 3948 AAGTAGCTTAAGTAACGCCATTTTGAAGGCATGGAATAATACATACTGAGAAATAGAGA 4007  
Db 3290 ACCGAGGTGCCTGTGTTTGAAGCCATCAGAGCAGAGATCTCCACACCCCAAGGCCAC 3349  
Qy 4008 AGTTCAGATCAAGGTTTAGGAACAGAGAGACAGCAGAAATATGGGCCCAACAGATATCTGT 4067  
Db 3350 ACTGTTATGCTGCGCCACAGGCTTCTACCCCGAACCACTGTGAGCTGAGCTGTGGGTGAA 3409

QY 4068 GGTAAAGCAGTTCTGCGCGCTCAGGCGCAAGAACAGTTGGAAACAGGAGAAATATGGCC-- 4125  
 Db |||||  
 QY 3410 TGGGAAGGAGGTGCAAGTGGGGTCAGACAGACCGCAGCCCTCAAGGAGCAGCCCGC 3469  
 Db |||||  
 QY 4126 -----CAAAACAGGATATCTGTGTAAAGCAGTTCTGCGCGCTCAGGCGCCCAAGAACAGA 4180  
 Db |||||  
 QY 3470 CCTCAATGATCTCCAGATATCTGCTAGCAGCGCCCTGAGGGTCTCGGCCACCTCTCGGCA 3529  
 Db |||||  
 QY 4181 TGGTCCCAGATGCGGTGCGCGCTCAGCAGTTCTAGAGAACCATCAGATGTTTCCAGG 4240  
 Db |||||  
 QY 3530 GAACCCCCCAACCACTTCGGCTGTCAAGTCCAGTTCTACGGGCTCTCGGAGAAATGACGA 3589  
 Db |||||  
 QY 4241 GTGCCCCAAGGACCTGAAATGACCTGTGCTTATTTGAACATAACCAATCAGTTGCTTTC 4300  
 Db |||||  
 QY 3590 GTGGACCAAGATAGGGCCAAACCCGTCAACAGATCGTCAGGCGCGAGGCTTGGGGTAG 3649  
 Db |||||  
 QY 4301 TCGCTTCTGTCGCGCGCTTCTGCTCCCGAGCTCAATAAAGAGAGCCCAACCCCTCAC 4360  
 Db |||||  
 QY 3650 AGCAGACTGTGGCTTCACTCCGAGTCTTACCAGCAAGGGTCTGCTGTGCCACCATCCT 3709  
 Db |||||  
 QY 4361 TGGCGCGCCAGTCTCCAGATAGACTGCTGCGCGGGTACCGGTGTTCTCAATAAACCC 4420  
 Db |||||  
 QY 3710 CTATGAGATCTTGTCTAGGGAAGGCCACCTTGTATGCCGTGCTGTGTCAGTGCCTCTGTC 3769  
 Db |||||  
 QY 4421 TCTTGCAAGTTGCATCGACTCGTGTGCTCGCTGTTCTTGGGAGGGTCTCTCTGAGTGAT 4480  
 Db |||||  
 QY 3770 GATGGCCATGG-----TCAAGAGAAAGGATTCAGAGGCTAGTGCACCTGCAAGCCA 3821  
 Db |||||  
 QY 4481 TGACTACCCGTCAGCGGGTCTTTTCAGTTTCTCCACCTACACAGGTCTCACTAACATTC 4540  
 Db |||||  
 QY 3822 AGCTTATCGATAAAATAAAGATTTTATTTAGTCTCCAGAAAGGGGGGAATGAAGAC 3881  
 Db |||||  
 QY 4541 CTGATGTGCCGAGGACTCGCTCAGCCCGT---TTTTGTTTATTAATAAATGCAAGAA 4597  
 Db |||||  
 QY 3882 CCCACCTGTAGGTTTGGCAAGCTAGCTTAAGTAAGCCCAATTTTGAAGGCATGGAATA 3941  
 Db |||||  
 QY 4598 CAGTGTTCCTTCAAGCCAGACTACATCTGACTCTCGGCTTTTATAAAGAAATGTTGAAG 4657  
 Db |||||  
 QY 3942 CATACTGAGATAGAGAAGTTCAGATCAAGTTAGGAACAGAGAGACAGATATGG 4001  
 Db |||||  
 QY 4658 GGCTCTGTGAGCTATCTGCCACACGACTTTTAAAGATTTTATGCTCTCGATGAGGA 4717  
 Db |||||  
 QY 4002 GCCAAACAGGATATCTGTGTAAAGAGTTCTGCGCGGCTCAGGSCCAAGAACAGATGG 4061  
 Db |||||  
 QY 4718 TTTAGTCAATCTATCTCTCTATTTTGTGCTGCTCTCGTATTTTAAATTTCTAGTTTG 4777  
 Db |||||  
 QY 4062 TCCCAGATGCGGTCCCGCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTG 4121  
 Db |||||  
 QY 4778 CACTCCCTTCTGAGAGCAGCGGATGCGAGATAGTTAATACTCTGAGGCGAGGCTTCT 4837  
 Db |||||  
 QY 4122 CCCCAGGACCTGAAATGACCTGTGCTTATTTG-AACTAACCAATCAGTTGCTTCT 4180  
 Db |||||  
 QY 4838 GTGAAAGGTGCTGCGGCTCAGTGTGAGATTTTGGCATATAAAGGGGTCTGCGCCCTGT 4897  
 Db |||||  
 QY 4181 CGCTTCTGTTCGGCGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACT 4240  
 Db |||||  
 QY 4898 GTACAGACAGATCGGAATCTAGATGTCATCTCAGAGTCCCGCGGGTTC----- 4947  
 Db |||||  
 QY 4241 CGGCGCGCAGTCCCTCCGATAGACTGCGTGGCGGGTACCGGTGTATCCAAATAAACCCCT 4300  
 Db |||||  
 QY 4948 -----GGGGCTCTGACTCAGGGCATCTTGTCTAGAGATCTCTACGCGGAGCA 4999  
 Db |||||  
 QY 4301 CTTGCAAGTTGCATCCGACTTGTGGTCTCGCTGTCTTGGAGGGTCTCTCTGAGTGAT 4360  
 Db |||||  
 QY 5000 TGTGGCCGGGTACCGAGC-----TCGAATTCGTAATCATGTCATAGCTGTTTCTGT 5053  
 Db |||||  
 QY 4361 TGAATACCGTCAAGCGGGTCTTTTCAGTATTCGTAATCATGTCATAGCTGTTTCTGT 4420  
 Db |||||  
 QY 5054 GTGAAATGTTTATCCGCTCACAATTCACAACAATACAGGCGGAGAGCATAAAGGTAA 5113  
 Db |||||  
 QY 4421 GTGAAATGTTTATCCGCTCACAATTCACAACAATACAGGCGGAGCATAAAGGTAA 4480  
 Db |||||

QY 5114 AGCCTGGGTGCTTAATGAGTGAGCTAACTACATTAATGCGTTCGCTCACTGCCCCC 5173  
 Db |||||  
 QY 4481 AGCCTGGGTGCTTAATGAGTGAGCTAACTACATTAATGCGTTCGCTCACTGCCCCC 4540  
 Db |||||  
 QY 5174 TTTTCAGTTCGGGAAACCTGTGTCAGTGTGATTAATGAAATCGGCGCAACGCGCGGGAG 5233  
 Db |||||  
 QY 4541 TTTTCAGTTCGGGAAACCTGTGTCAGTGTGATTAATGAAATCGGCGCAACGCGCGGGAG 4600  
 Db |||||  
 QY 5234 AGGCGGTTTGGCTATTTGGGCGCTTTCGGCTTCTCGCTCACTGACTCGCTCGGCTCGGT 5293  
 Db |||||  
 QY 4601 AGGCGGTTTGGCTATTTGGGCGCTTTCGGCTTCTCGCTCACTGACTCGCTCGGCTCGGT 4660  
 Db |||||  
 QY 5294 CCGTTCGGCTTGGCGGCGGCTATCAGCTCACTCAAAAGCGGTAATACCGGTTATCCACAGA 5353  
 Db |||||  
 QY 4661 CCGTTCGGCTTGGCGGCGGCTATCAGCTCACTCAAAAGCGGTAATACCGGTTATCCACAGA 4720  
 Db |||||  
 QY 5354 ATCAGGGGATTAACGCGAGGAAAGAAATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCG 5413  
 Db |||||  
 QY 4721 ATCAGGGGATTAACGCGAGGAAAGAAATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCG 4780  
 Db |||||  
 QY 5414 TAAAAAGCGCGTTCGCTGCGGCTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAA 5473  
 Db |||||  
 QY 4781 TAAAAAGCGCGGTTGCTGGCGCTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAA 4840  
 Db |||||  
 QY 5474 AAATCGAGCTCAAGTCAAGAGTGGGAAACCCGACAGGACTATAAAGATACACAGCGTT 5533  
 Db |||||  
 QY 4841 AAATCGAGCTCAAGTCAAGAGTGGGAAACCCGACAGGACTATAAAGATACACAGCGTT 4900  
 Db |||||  
 QY 5534 TCCCTCGGAAGCTCCCTGTCGCTTCCGAGTTCCTGTTCCGACCTTCCGCTTACCGGATACCT 5593  
 Db |||||  
 QY 4901 TCCCTCGGAAGCTCCCTGTCGCTTCCGAGGCTGCGCTTTTCCATAGCTCACGCTGAGGTATCT 4960  
 Db |||||  
 QY 5594 GTCGCGCTTCTCCCTTCCGGAAGCGTGGCGGCTTCTCATAGCTCACGCTGAGGTATCT 5653  
 Db |||||  
 QY 4961 GTCGCGCTTCTCCCTTCCGGAAGCGTGGCGGCTTTTCCATAGCTCACGCTGAGGTATCT 5020  
 Db |||||  
 QY 5654 CAGTTCGGTGTAGTGTGCTCCAAAGCTGGGCTGTGTGACAGAACCCCGCTTCAGGC 5713  
 Db |||||  
 QY 5021 CAGTTCGGTGTAGTGTGCTCCAAAGCTGGGCTGTGTGACAGAACCCCGCTTCAGGC 5080  
 Db |||||  
 QY 5714 CGACGCTGCGCTTATCCGGTAACTATCGTCTGAGTCCAAACCCGGTAAAGACAGACTT 5773  
 Db |||||  
 QY 5081 CGACGCTGCGCTTATCCGGTAACTATCGTCTGAGTCCAAACCCGGTAAAGACAGACTT 5140  
 Db |||||  
 QY 5774 ATCGGCACTGGCAGCAGCCTGCTAAAGGATTAAGCAGAGCGAGGTATGAGGCGGTCG 5833  
 Db |||||  
 QY 5141 ATCGGCACTGGCAGCAGCCTGCTAAAGGATTAAGCAGAGCGAGGTATGAGGCGGTCG 5200  
 Db |||||  
 QY 5834 TACAGAGTTCTTGAAGTGGTGGCTTAACTACGGCTACACTAGAAGGACAGTATTTGGTAT 5893  
 Db |||||  
 QY 5201 TACAGAGTTCTTGAAGTGGTGGCTTAACTACGGCTTACACTAGAAGGACAGTATTTGGTAT 5260  
 Db |||||  
 QY 5894 CTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTGTAGTCTTGTATCCGGCAA 5953  
 Db |||||  
 QY 5261 CTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTGTAGTCTTGTATCCGGCAA 5320  
 Db |||||  
 QY 5954 ACAAAACCCGCTGCTAGCGGTGGTTTTTGGTTTTCAGAGCAGAGATTAACGGCAGAAA 6013  
 Db |||||  
 QY 5321 ACAAAACCCGCTGCTAGCGGTGGTTTTTGGTTTTCAGAGCAGCAGATTAACGGCAGAAA 5380  
 Db |||||  
 QY 6014 AAAAGGATCTCAAGAAAGATCTTTTCATCTTTTCTACGGGCTCTGAGGCTCAGTGGAAACA 6073  
 Db |||||  
 QY 5381 AAAAGGATCTCAAGAAAGATCTTTTCATCTTTTCTACGGGCTCTGAGGCTCAGTGGAAACA 5440  
 Db |||||  
 QY 6074 AAATCACTGAAGGATTTTGGTTCATGAGATTAATCAAAAAGGATTTTCCACTAGATCT 6133  
 Db |||||  
 QY 5441 AAATCACTGAAGGATTTTGGTTCATGAGATTAATCAAAAAGGATTTTCCACTAGATCT 5500  
 Db |||||  
 QY 6134 TTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATAGTAACTTGGTCTGA 6193  
 Db |||||  
 QY 5501 TTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATAGTAACTTGGTCTGA 5560  
 Db |||||  
 QY 6194 CAGTTTACCAATGCTTTAATCACTAGTGAGGCGACCTATCTCAGCGATCTGTCTATTTCTCATC 6253  
 Db |||||

Db 5561 CAGTTACCAATGCTTAATCAGTGAGGACCTATCTCAGCGATCTGTCTATTTGGTTTCATC 5620  
Qy 6254 CATAGTTCCTGACTCCCGTGGTGTAGATAACTACGATACGAGGAGGCTTACCATTGCG 6313  
Db 5621 CATAGTTCCTGACTCCCGTGGTGTAGATAACTACGATACGAGGAGGCTTACCATTGCG 5680  
Qy 6314 CCCAGTTCCTGCAATGATACCGGAGACCCAGCTCACCAGGCTCAGATTTATCAGCAAT 6373  
Db 5681 CCCAGTTCCTGCAATGATACCGGAGACCCAGCTCACCAGGCTCAGATTTATCAGCAAT 5740  
Qy 6374 AAACAGCCAGCCGAGGAGGCGAGCGAGAGTGGTCTCAACTTTTATCCGCTCCAT 6433  
Db 5741 AAACAGCCAGCCGAGGAGGCGAGCGAGAGTGGTCTCAACTTTTATCCGCTCCAT 5800  
Qy 6434 CCAGTCTAATTAATTTGTCGCGGAGCTAGAGTAAGTAGTTTCGCGAGTTAATAGTTTGGC 6493  
Db 5801 CCAGTCTAATTAATTTGTCGCGGAGCTAGAGTAAGTAGTTTCGCGAGTTAATAGTTTGGC 5860  
Qy 6494 CAAGTTGTTGCCAATGCTACAGGC-TGCTGGTGTACGCTCGTGGTTGGTATGGCTTC 6552  
Db 5861 CAAGTTGTTGCCAATGCTACAGGCATCGTGGTGTACGCTCGTGGTTGGTATGGCTTC 5920  
Qy 6553 ATTCAGCTCCGGTTCCTCAACCATCAAGCGAGTTACATGATCCGCCATGTTGTGCAAAA 6612  
Db 5921 ATTCAGCTCCGGTTCCTCAACCATCAAGCGAGTTACATGATCCGCCATGTTGTGCAAAA 5980  
Qy 6613 AGCGTTAGCTTCCTTCGGTCTCCGATCGTTGTGAGAGTAAGTTGGCGCAGTGTATC 6672  
Db 5981 AGCGTTAGCTTCCTTCGGTCTCCGATCGTTGTGAGAGTAAGTTGGCGCAGTGTATC 6040  
Qy 6673 ACTCATGTTATGGCAGCACTGCATAATTTCTTACTGTCTATGTCATCCGTAAGATGCTT 6732  
Db 6041 ACTCATGTTATGGCAGCACTGCATAATTTCTTACTGTCTATGTCATCCGTAAGATGCTT 6100  
Qy 6733 TTCTGTACTGTTGAGTACTCAACCAAGTCAATCTGAGAACTAGTATGCGGCGACCGAG 6792  
Db 6101 TTCTGTACTGTTGAGTACTCAACCAAGTCAATCTGAGAACTAGTATGCGGCGACCGAG 6160  
Qy 6793 TTGTCTTTCGCGCGTCAATACGGGATAATACCGGCGACATACGAGAACTTTAAAGT 6852  
Db 6161 TTGTCTTTCGCGCGTCAATACGGGATAATACCGGCGACATACGAGAACTTTAAAGT 6220  
Qy 6853 GCTCATATTGGAACAGTTCTTCGGGCGGAAACTCTCAAGGATCTTACCGCTGTTGAG 6912  
Db 6221 GCTCATATTGGAACAGTTCTTCGGGCGGAAACTCTCAAGGATCTTACCGCTGTTGAG 6280  
Qy 6913 ATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTTCAGATCTTTTACTTTCAC 6972  
Db 6281 ATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTTCAGATCTTTTACTTTCAC 6340  
Qy 6973 CAGCGTTTCTGGTGAGCAAAAACAGGAAGCAAAATGCCGCAAAAAGGGAATAAGGC 7032  
Db 6341 CAGCGTTTCTGGTGAGCAAAAACAGGAAGCAAAATGCCGCAAAAAGGGAATAAGGC 6400  
Qy 7033 GACACGGAATGTTGCAATCTACTACTCTTCTCTTTTCAATTAATTAATGAACATTTATCA 7092  
Db 6401 GACACGGAATGTTGCAATCTACTACTCTTCTCTTTTCAATTAATTAATGAACATTTATCA 6460  
Qy 7093 GGGTTATTGTTCTATGAGCGGATACATATTGGAATGATTTAGAAAAATAACAAATAGG 7152  
Db 6461 GGGTTATTGTTCTATGAGCGGATACATATTGGAATGATTTAGAAAAATAACAAATAGG 6520  
Qy 7153 GGTTCGCGGACATTTCCCGAAAAGTGCCACCTGAGCTCTAAGAAACCATTTATTCAT 7212  
Db 6521 GGTTCGCGGACATTTCCCGAAAAGTGCCACCTGAGCTCTAAGAAACCATTTATTCAT 6580  
Qy 7213 GACATTAACCTATAAAAAATAGCGGTATCAGAGGCCCTTTTCGTCTCGCGGTTTCGGTGA 7272  
Db 6581 GACATTAACCTATAAAAAATAGCGGTATCAGAGGCCCTTTTCGTCTCGCGGTTTCGGTGA 6640  
Qy 7273 TGACGGTGAACCTCTGACATGACGCTCCGAGACCGTCAAGCTTGTCTGTGAAGC 7332

Db 6641 TGACGGTGAACCTCTGACATGACGATCCGAGAGCGGTCAAGCTTGTCTGTAAAGC 6700  
Qy 7333 GGATCCCGGAGCAGACAAGCCCGTCAGGCGCGCTCAGCGGTTGTTGCGGTTGTCGGG 7392  
Db 6701 GGATCCCGGAGCAGACAAGCCCGTCAGGCGCGCTCAGCGGTTGTTGCGGTTGTCGGG 6760  
Qy 7393 CTGCGTTAACTATCGGCGATCAGAGCAGATTTGTTACTGAGAGTGCACATATGCGGTTGA 7452  
Db 6761 CTGCGTTAACTATCGGCGATCAGAGCAGATTTGTTACTGAGAGTGCACATATGCGGTTGA 6820  
Qy 7453 AATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCATTCGCCATTCCAGGCT 7512  
Db 6821 AATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCATTCGCCATTCCAGGCT 6880  
Qy 7513 GCGCAACTGTTGGGAAGGCGGATCGGTGCGGCGCTCTTCGTATTACGCCAGCTGGCGAA 7572  
Db 6881 GCGCAACTGTTGGGAAGGCGGATCGGTGCGGCGCTCTTCGTATTACGCCAGCTGGCGAA 6940  
Qy 7573 AGGGGATGTCGTCGAAGGCGATTAAGTTGGGTAACCGCAGGTTTCCAGTCAAGAG 7632  
Db 6941 AGGGGATGTCGTCGAAGGCGATTAAGTTGGGTAACCGCAGGTTTCCAGTCAAGAG 7000  
Qy 7633 TTGTAAAACGAGCCAGTGCC 7654  
Db 7001 TTGTAAAACGAGCCAGTGCC 7022

RESULT 6  
US-10-471-065-20  
; Sequence 20, Application US/10471065  
; Publication No. US20040197855A1  
; GENERAL INFORMATION:  
; APPLICANT: Prof. Dr. Wiesmüller, Lisa  
; TITLE OF INVENTION: Test system for the determination of genotoxicities  
; FILE REFERENCE: P59532  
; CURRENT APPLICATION NUMBER: US/10/471,065  
; CURRENT FILING DATE: 2003-09-04  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 20  
; LENGTH: 9320  
; TYPE: DNA  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Description of the artificial sequence: Plasmid  
; OTHER INFORMATION: p5-Puro-CMV-(N'-EGFP)-CMV-Red-(EGFP-EU)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(1592)  
; OTHER INFORMATION: Retroviral vector p5NM  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (3374)..(3392)  
; OTHER INFORMATION: Retroviral vector p5NM  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (5527)..(9320)  
; OTHER INFORMATION: Retroviral vector p5NM  
; FEATURE:  
; NAME/KEY: gene  
; LOCATION: (1617)..(2216)  
; OTHER INFORMATION: Puromycin resistance gene from pRetroOn (Clontech,  
; OTHER INFORMATION: Palo Alto, CA, USA)  
; FEATURE:  
; NAME/KEY: promoter  
; LOCATION: (2267)..(2848)  
; OTHER INFORMATION: CMV promoter from pEGFP-N1 (Clontech, Palo Alto,  
; OTHER INFORMATION: CA, USA)  
; FEATURE:  
; NAME/KEY: gene  
; LOCATION: (2906)..(3348)  
; OTHER INFORMATION: N'-EGFP, derived from EGFP from pEGFP-N1  
; OTHER INFORMATION: (Clontech, Palo Alto, CA, USA)





Db 7400 CTGCGGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCA 7459  
Qy 5633 TAGCTACGCTGTAGGTATCTCAGTTCCGTTGAGTTCGTTCCCTCAAGCTGGGCTGTGT 5692  
Db 7460 ATGCTCAGCTGTAGGTATCTCAGTTCCGTTGAGTTCGTTCCCTCAAGCTGGGCTGTGT 7519  
Qy 5693 GCAGGAACCCCGCTTCAGCCGAGCCGCTGGCCTTATCCGGTAACTATCCTCTTGAGTC 5752  
Db 7520 GCAGGAACCCCGCTTCAGCCGAGCCGCTGGCCTTATCCGGTAACTATCCTCTTGAGTC 7579  
Qy 5753 CAACCCGTAAGACACAGCTTATCCCACTGGCAGCCAGCTGTAACAGATTTAGCAG 5812  
Db 7580 CAACCCGTAAGACACAGCTTATCCCACTGGCAGCCAGCTGTAACAGATTTAGCAG 7639  
Qy 5813 AGCAGGATNTAGCCGCTGTACAGAGTTCTTGAAGTGGTGGCTAACTACGCTACAC 5872  
Db 7640 AGCAGGATNTAGCCGCTGTACAGAGTTCTTGAAGTGGTGGCTAACTACGCTACAC 7699  
Qy 5873 TAGAAGNACAGTATTTGGTATCTGGCTCTCTCTGAGCCAGTTACCTTCGGAAAAAGAT 5932  
Db 7700 TAGAAGNACAGTATTTGGTATCTGGCTCTCTCTGAGCCAGTTACCTTCGGAAAAAGAT 7759  
Qy 5933 TGGTAGCTCTTGATCCGGCAAAACAAACCCAGCTGGTAGCGGTGGTTTTTGTGCA 5992  
Db 7760 TGGTAGCTCTTGATCCGGCAAAACAAACCCAGCTGGTAGCGGTGGTTTTTGTGCA 7819  
Qy 5993 GCAGCAGATTACGGCGAGAAAAAGGATCTCAAGAGATCCTTTGATCTTTCTACGGG 6052  
Db 7820 GCAGCAGATTACGGCGAGAAAAAGGATCTCAAGAGATCCTTTGATCTTTCTACGGG 7879  
Qy 6053 GTCTGAGCTCAGTGGAAACGAAAACTCAGTTAAGGATTTTGGTTCATGAGATTATCAAA 6112  
Db 7880 GTCTGAGCTCAGTGGAAACGAAAACTCAGTTAAGGATTTTGGTTCATGAGATTATCAAA 7939  
Qy 6113 AAGGATCTTCACTAGATCCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 6172  
Db 7940 AAGGATCTTCACTAGATCCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 7999  
Qy 6173 ATATGAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACTATCTCAGC 6232  
Db 8000 ATATGAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACTATCTCAGC 8059  
Qy 6233 GATCTGTCTATTTCGTTCACTCAGTTGCTGATCTCCCGTCTGTGAGATAACTACGAT 6292  
Db 8060 GATCTGTCTATTTCGTTCACTCAGTTGCTGATCTCCCGTCTGTGAGATAACTACGAT 8119  
Qy 6293 ACGGGAGGCTTACCATCTGGCCCGCAGTCTGCAATGATACCGGAGACCCACGCTCAC 6352  
Db 8120 ACGGGAGGCTTACCATCTGGCCCGCAGTCTGCAATGATACCGGAGACCCACGCTCAC 8179  
Qy 6353 GGCTCCAGATTATCAGCAATAAACCCAGCCAGCCGAGGCGGAGGCGAGAGTGTCTCC 6412  
Db 8180 GGCTCCAGATTATCAGCAATAAACCCAGCCAGCCGAGGCGGAGGCGAGAGTGTCTCC 8239  
Qy 6413 TGCAACTTTATCCGCTCCATCCAGTCTTAATTAATTTGTTCCGGGAAGCTAGAGTAAGTAG 6472  
Db 8240 TGCAACTTTATCCGCTCCATCCAGTCTTAATTAATTTGTTCCGGGAAGCTAGAGTAAGTAG 8299  
Qy 6473 TTCCGCGAGTTAATAGTTTGGCAACGTTTGGCAATGCTGTACAGGC-TGTTGGTGTACG 6531  
Db 8300 TTCCGCGAGTTAATAGTTTGGCAACGTTTGGCAATGCTGTACAGGCATCGTGGTGTACG 8359  
Qy 6532 CTGTCGTTGGTATGGCTTCATTCAGTCCGCTTCCCAAGCATCAAGGCGAGTTACATG 6591  
Db 8360 CTGTCGTTGGTATGGCTTCATTCAGTCCGCTTCCCAAGCATCAAGGCGAGTTACATG 8419  
Qy 6592 ATCCCCCATGTTGTGCAAAAAAGCGTTAGTCTCTCGTCTCCGATCGTTGTGAGAAG 6651  
Db 8420 ATCCCCCATGTTGTGCAAAAAAGCGTTAGTCTCTCGTCTCCGATCGTTGTGAGAAG 8479  
Qy 6652 TAAGTTGGCCGAGTGTATCAGTATGAGTTATGGCAGCACTGCAATATCTCTTACTGT 6711  
Db 8480 TAAGTTGGCCGAGTGTATCAGTATGAGTTATGGCAGCACTGCAATATCTCTTACTGT 8539

Qy 6712 CATCCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCAATTCTTGAGA 6771  
Db 8540 CATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCAATTCTTGAGA 8599  
Qy 6772 ATAGTGTATCGGCGAGCGAGTTGCTCTTTCGCCCGCTCAATACGGGTAATACCGGCC 6831  
Db 8600 ATAGTGTATCGGCGAGCGAGTTGCTCTTTCGCCCGCTCAATACGGGTAATACCGGCC 8659  
Qy 6832 ACATAGCAGAACTTTAAAGTGTCTCATATTGGAAGAGTTTCTTCGGGGCGAAACTCTC 6891  
Db 8660 ACATAGCAGAACTTTAAAGTGTCTCATATTGGAAGAGTTTCTTCGGGGCGAAACTCTC 8719  
Qy 6892 AAGGATCTTACCGCTGTGAGATCCAGTTGATGTAACCCACTCGTGACCCAACTGATC 6951  
Db 8720 AAGGATCTTACCGCTGTGAGATCCAGTTGATGTAACCCACTCGTGACCCAACTGATC 8779  
Qy 6952 TTACAGATCTTTTACTTTTACCAGGCTTTCGGGTGAGCAAAAACAGGAAGGCAAAATGC 7011  
Db 8780 TTACAGATCTTTTACTTTTACCAGGCTTTCGGGTGAGCAAAAACAGGAAGGCAAAATGC 8839  
Qy 7012 CGCAAAAAAGGGAATAAGGCGCACACGAAATGTTGAATACTCATCTCTCTCTTTTCA 7071  
Db 8840 CGCAAAAAAGGGAATAAGGCGCACACGAAATGTTGAATACTCATCTCTCTCTTTTCA 8899  
Qy 7072 ATATTATTGAAGCATTTTATCAGGGTTATTGTCTCATGAGCGGATACATATTGAAATGAT 7131  
Db 8900 ATATTATTGAAGCATTTTATCAGGGTTATTGTCTCATGAGCGGATACATATTGAAATGAT 8959  
Qy 7132 TTAGAAAAATAACAAATAGGGTTCGCGCACATTTCCCGGAAAAAGTGCCACCTGACGT 7191  
Db 8960 TTAGAAAAATAACAAATAGGGTTCGCGCACATTTCCCGGAAAAAGTGCCACCTGACGT 9019  
Qy 7192 CTAGAAACCATATTATCATGACATTAACCTATAAAATAGGCGTATCACGAGGCCCTT 7251  
Db 9020 CTAGAAACCATATTATCATGACATTAACCTATAAAATAGGCGTATCACGAGGCCCTT 9079  
Qy 7252 TCCTCTCGCGGTTTCGGTGATGACGGTGAAACCTCTGACACATGACAGTCTCCGGAGAC 7311  
Db 9080 TCCTCTCGCGGTTTCGGTGATGACGGTGAAACCTCTGACACATGACAGTCTCCGGAGAC 9139  
Qy 7312 GGTACAGCTTGTCTGTAAGCGGATGCGGAGCAGACAGCCCGTTCAGGGCGGTGACG 7371  
Db 9140 GGTACAGCTTGTCTGTAAGCGGATGCGGAGCAGACAGCCCGTTCAGGGCGGTGACG 9199  
Qy 7372 GGGTGTGGCGGTTGTCGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAG 7431  
Db 9200 GGGTGTGGCGGTTGTCGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAG 9259  
Qy 7432 AGTGACCATATGCGGTGTGAAATACCGCAGATGCGTAAGGAGAAAAATACCGCATCAG 7491  
Db 9260 AGTGACCATATGCGGTGTGAAATACCGCAGATGCGTAAGGAGAAAAATACCGCATCAG 9319  
Qy 7492 G 7492  
Db 9320 G 9320

## RESULT 7

US-09-982-223A-1  
; Sequence 1, Application US/09982223A  
; Publication No. US20030175972A1  
; GENERAL INFORMATION:  
; APPLICANT: Daley, George Q.  
; APPLICANT: Koh, Eugene Y.  
; TITLE OF INVENTION: EXPRESSION VECTORS AND USES THEREOF  
; FILE REFERENCE: 13086-002001  
; CURRENT APPLICATION NUMBER: US/09/982,223A  
; CURRENT FILING DATE: 2001-10-18  
; PRIOR APPLICATION NUMBER: 60/241,879  
; PRIOR FILING DATE: 2000-10-20  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Fast-Seq for Windows Version 4.0





Db 1977 GGGACACTTCACATGAGCGAAATAATACATCGTCACTGGGACATGTTGACAGATCCATGC 2036  
Qy 2587 TGGTACACAGAGAGCGAGTAAAGCTCCAAAGCTGCTGATCTACTGAGACATCCACCCGG 2646  
Db 2037 ACCTAACTCGCAAGCCGACTGATGC-----CTTCTGAACAATGGAAGGCATTAATG 2089  
Qy 2647 CACACTGCTGTGCAAGCAGATTCAGCGGTAGCGGTAGCGGTACCGACTTCACCTTCACC 2706  
Db 2090 CCGTAAGCCGTGGCGGTCTGGTACCGGTGGGTGAAGACCAGAAACAGCACCTCGATC--- 2146  
Qy 2707 ATCAGACGCTCCAGCCAGAGGACATCGCCACCTACTACTGCCAGCAATATAGCCTCTAT 2766  
Db 2147 -TGAGCGCGATATTGCCACGGTTTCAACCGCTGTATGCGGAGATCGATCCCGTC--- 2202  
Qy 2767 CGGTGTTCCGCCCAAGGAGCAAGGTGGAATCAAAAGAGGTGCTCAGGATCGGGTGA 2826  
Db 2203 -----GTTTTCACCGTCTGACTGGGAAACCTCGCGTTA 2239  
Qy 2827 TCCGGCTCTGGTCTCAGGATCGGAGTCCAACTGGTGGAGAGCGGTGGAGGTGTTGTG 2886  
Db 2240 CCAACTTAATGCGCTTGGAGGACATCCCTCTTTCGCCAGCTGGCGTAAATAGCGAAAGG 2299  
Qy 2887 CAACCTGCGCTCCCTCGCTCTGCTCGCTCCGCTCGCATCTGCTTCCGATTTTCCACACATAT 2946  
Db 2300 CCGCACCGATCGCCCTTCCCAAAGTTCGCGAGCTGAAATGGGGAATGGCGTTTGCC 2359  
Qy 2947 TGAATGAGTGGGTGAGACAGGCACTGGAAAGGTCTTGAAGTGAATGGAGAAATCAT 3006  
Db 2360 TGGTTTCGGCACAGAAAGCGGTGCGGAAAGCTGGCTGGAGTGGAT----- 2407  
Qy 3007 CCAGTAGCAGTAACTATGCGCGTCTTAAGAGTATGTTTACAAATATCGGA 3066  
Db 2408 -CTTCTGAGGCGGATACTGCTGCTGCTCCCTCAAACTGGCAGATGACCGGTTTACGATG 2465  
Qy 3067 GACAAAGCAAGAAACACATGTTCTGCAAAATGACAGCCTGAGACCGAAGACACCGGG 3126  
Db 2466 CGCCATCTACCAAGCTGACCTATCCATACGTCATCCGCGTGTGTTTCCACGG 2525  
Qy 3127 GTCTATTTTGTCAAGCCTTTACTTCCGCTTCCCTGCTTGTGTTTGTATTTGGGCGCAAGGG 3186  
Db 2526 AGAATCCGACGGTGTGTTACTCGCTCACATTTTAATGTTGATGAAGCTGCTACAGGA 2585  
Qy 3187 ACCCGGTACCGTCTCAGTGTCTAAGCCACAGAGCGAGCGCGCGACACCAACA 3246  
Db 2586 GGCAGACGCGAATTAATTTTGTGAGCGGTAACTCGCGCTTTCATCTG----- 2633  
Qy 3247 CCGCGCCCAACATCGCTGCGAGCCCTGCTGCTGCGCCAGAGGGCGCTCGGCCAGCG 3306  
Db 2634 -----TGGTGCACGGCGCTGGGTTCGGTTCAGGCGAAGACAGTCTG 2674  
Qy 3307 GCGGGGGCGAGTGACACAGAGGGGGCTGGACTTTGCGCTGGATCCCAACTCTGCTAC 3366  
Db 2675 TTTGGCTCTTAATTTGAGCTCGAGCGCATATCTAGCGCGCGA-----GNAACCGCCTC 2730  
Qy 3367 CTGCTGAGTGAATCTCTTCACTATGCTGTCTATCTCACTGCTGTTTCTGAGAGTG 3426  
Db 2731 CGCGTGTGTGCTGCTGCTGAGTGAACGGAGTATCT----- 2768  
Qy 3427 AAGTTCAGAGGCGAGCGCCCGCTGCTGCTGCGCCAGAGGGCGCAGAACCACTCTATAC 3486  
Db 2769 -----TGAAGATCAAGATATGCGGATGAGCGGATTCGAGCGAANAAC 2814  
Qy 3487 GAGCTCAATCTAGGACGAAGAGAGGATGACATGTTTGGACAAGAGACGTGGCGCGGAC 3546  
Db 2815 GGTCTGCGCTCGGGAACGCGGAATTTGAATTTATGGCCACACAGAGTGGCGCGCGAC 2874  
Qy 3547 CTTGAGATGGGGGGAAGCCGAGAGGAAGAAACCTCTAGGAAGCGCTGTACAATGAACCTG 3606  
Db 2875 TTCCAGTT-----CAACATCAGCGCTACAGTCAACAGCAACTGATGGAACACCGCCAT 2928  
Qy 3607 CAGAAAGATAGATGGCGGAGCCCTACAGTCAAGATGGGATGGAAGGCGAGCGCGGAGG 3666  
Db 2929 CGCCATCTGCTGCGAGGAAAGAACCCGACATGGGCTGTTATACGACGGTTCATATGGGG 2988

Qy 3667 GGAAGGGGACAGATGCGCTTTACCGGGTCTCAGTACAGCCACCAAGGACACCTACGAC 3726  
Db 2989 ATTGGTGGGACGACTCTCTGAGCCCGTCAGTATCGCGGAATTCACAGTGGCGCGCGT 3048  
Qy 3727 GCCCTTCACATGCAAGGCCCTGCCCCCTCGCTAACTCGACGGCGCGCGGATCCGATTAG 3786  
Db 3049 CGCTACCATTAACCAAGTTGGTCTGGTGTCAAAAATAATAATAACCGGCGAGCCATGCTG 3108  
Qy 3787 TCCAATTTGTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTTTGTATCAACAATATC 3846  
Db 3109 CCGTATTTTCGCGTAAGGAAATCCATTATGT-----ACTATTTA 3147  
Qy 3847 ACCAGCTGAAGCTTATAGAGTACGAGCCATAGATAAAATAAAGATTTTATTTTGTCTCC 3906  
Db 3148 AACTCGAGCGCGCCCGCCAGCACAGTGGTGCAGATAAAATAAAGATTTTATTTTGTCTCC 3207  
Qy 3907 AGAAAAAGGGGGAATCAAGAGACCCCACTGTAGTTTGGCAAGCTAGCTTAAGTAAGCC 3966  
Db 3208 AGAAAAAGGGGGAATGAAGACCCCACTGTAGTTTGGCAAGCTAGCTTAAGTAAGCC 3267  
Qy 3967 CATTTTCAAGGCAATGG-AAAAATACAACTGAGAAATAGAGAATTCAGATCAAGGTTAG 4025  
Db 3268 CATTTTCAAGGCAATGAAAAATACATAACTGAGAAATAGAGAATTCAGATCAAGGTCAG 3327  
Qy 4026 GAACAGA-GAGACAGCAGATATGGGCCAAACAGGATATCTGTGTAGAGAGTTCCTGCC 4084  
Db 3328 GAAAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGTAGAGAGTTCCTGCC 3387  
Qy 4085 CC-GCTCAGGCCCAAGAACAGTTCGGAACAGGAGATATGGGCCAAACAGGATATCTGTGG 4143  
Db 3388 CCGCTCAGGCGCAAGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGG 3447  
Qy 4144 TAACAGTTCCTGCCCCGGCTCAGGGCCAGAAACAGATGTTCCCGAGATCGGTCGCCGCC 4203  
Db 3448 TAACAGTTCCTGCCCCGGCTCAGGGCCAGAAACAGATGTTCCCGAGATCGGTCGCCGCC 3507  
Qy 4204 CTACAGCTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCAAGAACCTGAAATGAC 4263  
Db 3508 CTACAGCTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCAAGAACCTGAAATGAC 3567  
Qy 4264 CTTGTGCTTATTTGAATTAACCAATCAGTTCGCTTCTCGCTTCTGTTTCGCGCGCTCTG 4323  
Db 3568 CTTGTGCTTATTTGAATTAACCAATCAGTTCGCTTCTCGCTTCTGTTTCGCGCGCTCTG 3627  
Qy 4324 CTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGCGCGCCAGTCTCTCGATAG 4383  
Db 3628 CTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGCGCGCCAGTCTCTCGATAG 3687  
Qy 4384 ACTGCTGCGCGGGTACCGGTGTTCTCAATAAACCCCTTTGTCAGTTGTCATCCGACTCGT 4443  
Db 3688 ACTGAGTGGCGGGTACCGGTGTTCTCAATAAACCCCTTTGTCAGTTGTCATCCGACTTGT 3747  
Qy 4444 GGTCTGCTGTTCTTGGAGGGTCT-CTCTGAGTGTGATGACTACCGCTCAGCGGGTCT 4502  
Db 3748 GGTCTGCTGTTCTTGGAGGGTCTCTCTGAGTGTGATGACTACCGCTCAGCGGG--- 3804  
Qy 4503 TTCAGTTTCTCCACCTACACAGGTCTCACTAACTCTGATGTCGCGCAGGAGCTCCG 4562  
Db 3805 ----- 3804  
Qy 4563 TCAGCCCCGTTTGTGTTTATAATAAATGCAAGAACAGTGTTCCTTCAAGCCAGACTAC 4622  
Db 3805 ----- 3804  
Qy 4623 ATCTGACTCTCGGCTTTATAAAGAAATGTTGAAGGGCTCTGTGACTATCTGCCACAG 4682  
Db 3805 ----- 3804  
Qy 4683 ACTTTTAAAGATTTTATGCTCTCGGATGAGGATTTAGTCAATCTATCTCTGCTATT 4742  
Db 3805 ----- 3804

Qy	4743	TTGCTGGCTTCTCCGTATNTTTAAATTTCTTAGTTTGGCACTCCCTTCTCTGAGACACGGCGA	4804
Db	3805	-----	3804
Qy	4803	TTGCAGAGTAGTTAATACTCTGAGGGCAGGCTTCTGTGAAGAAGTTCCTTGGGCTCAGTG	4862
Db	3805	-----	3804
Qy	4863	TGAGATTTGCCATAAAAAAGGGTCTCTGCCCTGTGTACAGACAGATCGGAATCTAGAGT	4922
Db	3805	-----	3804
Qy	4923	GCATACTCAGAGTCCCOCGGTTCCTGGGGCTCTGATCTCAGGCGATCTTTTGCTTAGAGAT	4982
Db	3805	-----	3804
Qy	4983	CCTCTACGCCGAGCATCGTGGCCGGTACCGAGCTCGAATTCTGTAATCATGGTCAATAG	5042
Db	3805	-----	3804
Qy	5043	CTGTTTCTGTGTAATTTGTTATTCGCTCACAATTCACAACAATCAGCGCGGAAGC	5102
Db	3805	-----	3804
Qy	5103	ATAAAGTTAAAGCTCGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGC	5162
Db	3805	-----	3804
Qy	5163	TCACTGCCCGTTTCCAGTCGGGNAACCTGTGTGCGCAGCTGCATTAATGAATCGGCGAA	5222
Db	3805	-----GTCTTTTCAATTCGTCAATTAATGAATCGGCGAA	3835
Qy	5223	CGCGCGGGAGAGCGGTTTTCGTTATTTGGGGCTCTTCGCTTCTCGCTCACTGACTCG	5282
Db	3836	CGCGCGGGAGAGCGGTTTTCGTTATTTGGGGCTCTTCGCGCTTCTCGCTCACTGACTCG	3895
Qy	5283	CTCGCTCGGTTCGGTTCGGTCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAAATACGG	5342
Db	3896	CTCGCTCGGTTCGGTTCGGTCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAAATACGG	3955
Qy	5343	TTATCCACAGAAATCAGGGGATACCGCAGGAAAGAAATGTGAGCAAAAGGCCAGCAAAG	5402
Db	3956	TTATCCACAGAAATCAGGGGATACCGCAGGAAAGAAATGTGAGCAAAAGGCCAGCAAAG	4015
Qy	5403	GCCAGGAACGTTAAAAAGCGCGTTCCTGCGCTTTTTTCCATAGGCTCCGCCCCCTGAC	5462
Db	4016	GCCAGGAACGTTAAAAAGCGCGGTTCTCGCGCTTTTTTCCATAGGCTCCGCCCCCTGAC	4075
Qy	5463	GAGCATCAAAAAATCGACGCTCAAGTFCAGAGGTGGGAAACCCGACAGACTATATAAGA	5522
Db	4076	GAGCATCAAAAAATCGACGCTCAAGTFCAGAGGTGGGAAACCCGACAGACTATATAAGA	4135
Qy	5523	TACAGCGGTTTCCCTCGAAGCTCCCTCGTGGCTCTCTCTGTTTCGACCTTCGCGCTT	5582
Db	4136	TACAGCGGTTTCCCTCGAAGCTCCCTCGTGGCTCTCTCTGTTTCGACCTTCGCGCTT	4195
Qy	5583	ACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGC	5642
Db	4196	ACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGC	4255
Qy	5643	TGTAGGTATCTCAGTTCCGTTGAGTTCGTTTCGCTCCAAGCTGGGCTGTGTGACGAAACC	5702
Db	4256	TGTAGGTATCTCAGTTCCGTTGAGTTCGTTTCGCTCCAAGCTGGGCTGTGTGACGAAACC	4315
Qy	5703	CCGTTTCAGCCCGACCGCTGGCCCTTATCCGGTAACTATCGTCTTGAGTCCAAACCGGTA	5762
Db	4316	CCGTTTCAGCCCGACCGCTGGCCCTTATCCGGTAACTATCGTCTTGAGTCCAAACCGGTA	4375
Qy	5763	AGACACACTTATTCGCCACTTCGGCAGCAGCCACTTGGTAAACAGGATTAAGCAGAGCGGTAT	5822
Db	4376	AGACACACTTATTCGCCACTTCGGCAGCAGCCACTTGGTAAACAGGATTAAGCAGAGCGGTAT	4435
Qy	5823	GTAGGCGGTGCTACAGAGTTCTTTGAAGTGTGGTGCCTAACTTACGGCTACACTAGAAGACA	5882

4436	Db		GTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCCTAACTACGGCTACACTAGAGGACA	4495
5883	Qy		GTATTTGGTATCTGGCCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGCTAGCTCT	5942
4496	Db		GTATTTGGTATCTGGCCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGCTAGCTCT	4555
5943	Qy		TGATCCGGCAAAACAAACCCACCGCTGGTGTAGCGGTGGTTTTTTTGTGTTTGAAGCAGCAGATT	6002
4556	Db		TGATCCGGCAAAACAAACCCACCGCTGGTGTAGCGGTGGTTTTTTTGTGTTTGAAGCAGCAGATT	4615
6003	Qy		ACGCGCAGAAAAAAGATCTCAAGAAAGTCTTTTGATCTTTTCTACGCGGTCTGACGCT	6062
4616	Db		ACGCGCAGAAAAAAGATCTCAAGAAAGTCTTTTGATCTTTTCTACGCGGTCTGACGCT	4675
6063	Qy		CAGTGAACGAAACACTCAGTTAAGGATTTTGGTTCATGAGATTATCAAAAAGATCTTTC	6122
4676	Db		CAGTGAACGAAACACTCAGTTAAGGATTTTGGTTCATGAGATTATCAAAAAGATCTTTC	4735
6123	Qy		ACCTAGATCTTTTAAATTAATAATCAAGTTTAAATCAATCTAAAGTATATATAGTAA	6182
4736	Db		ACCTAGATCTTTT-----GGCGCGCGCGCAATCAATCTAAAGTATATATAGTAA	4788
6183	Qy		ACTTGGTCTGACAGTTTACCAATGCTTTAATCAGTGAAGCACTTATCTCAGCGATCTGCTTA	6242
4789	Db		ACTTGGTCTGACAGTTTACCAATGCTTTAATCAGTGAAGCACTTATCTCAGCGATCTGCTTA	4848
6243	Qy		TTTTCGTTTCAATCAGTATGCTGACTCCCGTCTGTAGATAAATACTAGGAGGCG	6302
4849	Db		TTTTCGTTTCAATCAGTATGCTGACTCCCGTCTGTAGATAAATACTAGGAGGCG	4908
6303	Qy		TTACCACTCTGGCCCCAGTCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGAT	6362
4909	Db		TTACCACTCTGGCCCCAGTCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGAT	4968
6363	Qy		TTATCAGCAATAAACACAGCCAGCCGAAGGCGCGAGCAGAGTGGTCTTGCAACTTTTA	6422
4969	Db		TTATCAGCAATAAACACAGCCAGCCGAAGGCGCGAGCAGAGTGGTCTTGCAACTTTTA	5028
6423	Qy		TCCGCTCCATCCAGTCTATTAAATGTTTCCGCGGAGCTAGAGTAGTAGTTGCGCAGTT	6482
5029	Db		TCCGCTCCATCCAGTCTATTAAATGTTTCCGCGGAGCTAGAGTAGTAGTTGCGCAGTT	5088
6483	Qy		AATAGTTTGGCAACGTTGTGTCATTGCTACAGGG-TCGTGGTGTACAGCTCGTCTGTTT	6541
5089	Db		AATAGTTTGGCAACGTTGTGTCATTGCTACAGGGCATGCTGCTGTCAGCTCGTCTGTTT	5148
6542	Qy		GGTATGGCTTCATTTACGCTCCGGTTCCCAACGATCAAGGCGAGTTTACATGATCCCCATG	6601
5149	Db		GGTATGGCTTCATTTACGCTCCGGTTCCCAACGATCAAGGCGAGTTTACATGATCCCCATG	5208
6602	Qy		TTGTGCAAAAAGCGGTTAGTCTCTTCGGTCTCCCGATGTTGTGTCAGAAAGTAGTTGGCC	6661
5209	Db		TTGTGCAAAAAGCGGTTAGTCTCTTCGGTCTCCCGATGTTGTGTCAGAAAGTAGTTGGCC	5268
6662	Qy		GCAGTGTATCATCTCATGTTATGCGCAGCACTGCATAATTTCTTACTGTCATGCCATCC	6721
5269	Db		GCAGTGTATCATCTCATGTTATGCGCAGCACTGCATAATTTCTTACTGTCATGCCATCC	5328
6722	Qy		GTAAAGATGCTTTTCTGTGATCTGGTGAAGTACTCAACCAAGTCAATCTGAGAATAGTGATG	6781
5329	Db		GTAAAGATGCTTTTCTGTGATCTGGTGAAGTACTCAACCAAGTCAATCTGAGAATAGTGATG	5388
6782	Qy		CGCGCAGCGAGTTGCTCTTGGCCGGGCTCAATACGGGATTAATACCGCGCCACATAGCAGA	6841
5389	Db		CGCGCAGCGAGTTGCTCTTGGCCGGGCTCAACACGGGATTAATACCGCGCCACATAGCAGA	5448
6842	Qy		ACTTTAAAAGTGTCTCATCTGAAACGTTCTTGGGGCGGAAAACTCTCAAGGATCTTTA	6901
5449	Db		ACTTTAAAAGTGTCTCATCTGAAACGTTCTTGGGGCGGAAAACTCTCAAGGATCTTTA	5508
6902	Qy		CCGCTCTTGAGATCCAGTTTCGATGTAAACCACTCGCTGCAACCAACTGATCTTCAGCATCT	6961



1980 CTTCTCTGTTGACCCCGCTCGATCCTCCTCTTATCCAGCCCTCACTCCTCTCTAGGC 2039  
1342 CTTCTCTGTTGACCCCGCTCGATCCTCCTCTTATCCAGCCCTCACTCCTCTCTAGGC 1401  
2040 GCGCCCATATGGCCATATAGATCTTATATAGGGGCAACCCCGCCCTTGTAAACTTCCCT 2099  
1402 GCGGAGT-----CT 1411  
2100 GACCTGACATGACAGAGTTACTAACAGCCCTCTCTCCAAGCTCACTTACAGGTTCT 2159  
1412 CTGAGGAGTTAAACGAGTTTAAACGAGCGCGCCGCGCAAGCT----- 1456  
2160 ACTTAGTCCAGCACGAAGTCTGAGACCTCTGGCGCACCTTACCAAGAACAACATGGACC 2219  
1457 -----TGAGGAAATCCGCCCCCTCTCCTCCCGCCCTTAACAGTTACTGGCGAA 1506  
2220 GACCGGTGTA CTTCACTTACCGAGTCGGGACACAGTGTGGTCCCGACACAG 2279  
1507 GCGCTTGAATAAGCGCGTGTGCTTGTCTATATGTTATTTTCCACCATATTGCGCT 1566  
2280 CTAAGNACCTAGAACCTCGCTGGAAAGACCTTACACAGTCTCTGTGACACCCCGCCACCG 2339  
1567 CTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTGTGACGAGCATTTCTTAGGG 1626  
2340 CCTCAAAGTAGACGGCATCGAGCTTGGATACACGCCGCCAGGTGAAGGCTGCCGACC 2399  
1627 GTCTTTCCCTCTCGCCAAAGGATGCAAGTCTGTGATGTGTGTAAGGAAGAGTTTC 1686  
2400 CCGGGGTGGACATCTCTAGACTCCATGGATGGAGCTGTATCATCTCTCTTTGTT 2459  
1687 CTCTG--GAAGCTTTGAAGACAAACAACGTCTGTGCGACCTTGTGCGCCCTTTGCG 1733  
2460 AGCAACAGTACAGGTGTCTCACTCCGACATCAGCTGACCCAGAGCCGCAAGCGCTGAG 2519  
1734 -----AGGACGGGAACCCCGCCCTGCGGACAGGTGCTCTGCGGCCAAAGGCCACGTG 1788  
2520 CGCAGCGTGGTGACAGAGTGACCATCACCTGTGAAGCCAGTCAGGATGTGGTACTTC 2579  
1789 TATAGATACACTGTCGAAGGGGACAAACCCAGTGCACCGTGTGATGTGATGATGTTG 1848  
2580 TGTAGTCTGTACAGCAGAGCCAGGTAAAGGCTCCAAAGCTGCTGATCTACTGGACATC 2639  
1849 TGGAAAGATCA----- 1860  
2640 CACCGGACACATGTGTGTCCAGAGCATTCAGCGTAGCGGTAGCGGTACCGACTTCAC 2699  
1861 ----- 1860  
2700 CTTTACCATCAGCGCTCCAGCCAGAGGACATCGCCACCTACTACTGCCAGCAATATAG 2759  
1861 -----ATGGCT 1867  
2760 CTTCTATCGGTGTGGCCAAAGGACCAAGGTGGAATCAAAACGAGGTGGCTCAGGATC 2819  
1868 CTTCTCAAGCGTATTCAACAAGGGCTGAAGATG-----CCAGAGGTATCCCAATT 1920  
2820 GGTGTATCCGCTCTGTGTGGCTCAGGATCGGAGTTCAACTGGTGGAGAGCGGTGGAGG 2879  
1921 GTATGGGATCTGATCTGGGGCCCTC-----GGTGACATGCTTTTACATGTGTTAGT 1971  
2880 TGTGTGCAACCTGGCGCGTCTCGCTGTCTGCTCGCATCTGGCTTCGATTTTCAC 2939  
1972 CGAGGTTAAACAGTCTTAGGC-----CCCCGAAACACCGGGACGTGGTTTCTTTGAA 2028  
2940 CACATATTGGATGAGTTGGGTGAGACAGGCACCTGGAAAGGTCTTTGATGTGATTTGAGA 2999  
2029 AACACGATGATAATATGGCCACAAACCAAGGGGAGAGCTGTTACCGGGGTGGTCCCA 2088  
3000 AATTCAATCCAGATAGCAGTTACGATTAACTATGCGCGTCTCTAAAGGATAGATTTACAA 3059  
2089 TCTGTGTCGAGCTG----- 2103

3060 ATCGGAGACAACGCGCAAGAACACATTTGTTCTGCAATGGACAGCTTGAGACCCGGAAGA 3119  
2104 -----ACGGCGACGTGAACGCGCACAGTTTCAGGTGTCGCGGAGGG 2146  
3120 CACCGGGGTCTATTTTGTGCAAGCTTTTACTTGGCTTCCCTGTTGCTTTATTTGGGG 3179  
2147 CAGGGCGCATGTCACCTAGCGCAAGCTGACCCCTGAAGTTTCTGTGCACCAACCGCAAGCT 2206  
3180 CCAAGGGACCCCGGTCCCGTCTCAGTCTCTAAGCCCAACGACGCGCAGCGCCGACCC 3239  
2207 GCGGTGCTGCGCCACCCCT-----CGTGACC 2234  
3240 ACCAAACCGCGCGCCACCATCGCTCGCAGCCCTCTGCTGCGCCCGCAGAGCGGCTCG 3299  
2235 ACCCTGACCTAGCGGTGCGATGCTT--CAGCGCTACCCCGACACATGAGCAGCAG 2292  
3300 GCCAGCGGGGGGGCGGAGTGCAACAGAGGGGTGGAATTTGCGCTTGATTCCTAAACT 3359  
2293 AC-----TTCTTCAAGTCCGCCATGCCGGAAG 2320  
3360 CTGCTACCTGCTGGATGGAATCCTCTTCTATCTATGCTGTCTATCTCACTGCTCTTCT 3419  
2321 CTACGTCAGGAGCGCACCATCTCTTCAAGGACGACGCAACTCAAGACCCGCGCGCA 2380  
3420 GAGAGTGAAGTTTACGACGAGCGCAGAGCCCGCTACGACGAGGCGCAGAACGAGCT 3479  
2381 GGTGAAGTTGAGGGGGACACCTCTGTGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAA 2440  
3480 CTATTAACGAGCTCAATCTTAGGACGAAGAGAGTAGTATGATTTTGGACAAGAGAGTGG 3539  
2441 GGAGACGCGCAACATCTGGGGCAAGCTGGAGTACAACACTACAACAGCACCAAGTCTTA 2500  
3540 CCGGACCTTGAGTGGGGGGAAGCCGAGAGGAAGAACCTCAGGAAGGCTCTGTCAA 3599  
2501 TATCATGCGCGCAAGAGCAAGCGCATCAAGCGCAACTTCAAGATCCCGCCACAACAT 2560  
3600 TGAAGTCAGAAAGATAAGATGGCGGAGCGCTTACAGTAGATGGGATGAAGAGGCGAGCG 3659  
2561 CGAGGACGCGAGCTGTGAGCTCGCGACCACTACAGCAGAACACCCCATCGCGCAGCG 2620  
3660 CCGGAGGGCAAGGGGCAAGATGGCTTTTACAGGGTCTCAGTACAGCCACCAAGACAC 3719  
2621 CCGCGTGTCTGCTCGCGACCACTACTCTGAGCACCAGTCCGCTCGCAGCAAGACCC 2680  
3720 CTACGACCGCTTACATGCGAGCGCTCGCCCTCGTAACTCGACGCGCGCGGATCC 3779  
2681 CAACGAGAGCGCATCATGCTGCTGGAGTCTGTGACCCCGCGCGGATCACTCA 2740  
3780 GGATTTAGTCCAAATTTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTGTACTCAA 3839  
2741 CGGCATGGACGAGCTGTACAGTAAGTCGACTGCGAGCCNAGCTT----- 2785  
3840 CAATATCACCAGCTGAAGCTTATAGAGTACGAGCATAGATAAATAAAGATTTTATTT 3899  
2786 -----ATCGATAAATAAAGATTTTATTT 2810  
3900 AGTCTCAGAAAAAGGGGGATGAAGACCCCACTGTAGGTTTGGCAGTAGCTTAA 3959  
2811 AGTCTCAGAAAAAGGGGGATGAAGACCCCACTGTAGGTTTGGCAGTAGCTTAA 2870  
3960 GTAACGCGCTTTTGAAGGATGGAATAATACTAGATAAGAGTTCAGATCAA 4019  
2871 GTAACGCGCTTTTGAAGGATGGAATAATACTAGATAAGAGTTCAGATCAA 2930  
4020 GGTTAGGAAACAGAGAGACAGCAGATAATGGGCGCAACAGGATATCTGTGGTAAGCAGTTC 4079  
2931 GGTTAGGAAACAGAGAGACAGC----- 2951  
4080 CTGCGCCCGCTCAGGGCCAGAGACAGTTGGAACAGAGATATGGGCCCAACAGGATATCT 4139  
2952 -----AGAATATGGGCCCAACAGGATATCT 2976  
4140 GTGGTAAGCAGTTTCTGCGCCCGCTCAGGGCCAGAAACAGATGGTCCCCCAGATGGGTCC 4199







QY 6358 CAGATTATCAGCAATAAACACGACGACGCGAAGGCGGAGCGCAGAAAGTGGTCTCGCAA 6417  
 DB 4684 CAGATTATCAGCAATAAACACGACGACGCGAAGGCGGAGCGCAGAAAGTGGTCTCGCAA 4743  
 QY 6418 CTTTATCCGCTCCATCCAGTCTATTAAATTTGTCGCGGAAGCTAGAGTAAGTAGTTCG 6477  
 DB 4744 CTTTATCCGCTCCATCCAGTCTATTAAATTTGTCGCGGAAGCTAGAGTAAGTAGTTCG 4803  
 QY 6478 CAGTTAATAGTTGCGCAAGTGGTGGCAATGCTACAGCG-TCGTGGTGTACGCTCGT 6536  
 DB 4804 CAGTTAATAGTTGCGCAAGTGGTGGCAATGCTACAGCGCATCGTGGTGTACGCTCGT 4863  
 QY 6537 CGTTTGGTATGGCTTCATTACGCTCCGTTTCCCAACGATCAAGCGGAGTTACATGATCCC 6596  
 DB 4864 CGTTTGGTATGGCTTCATTACGCTCCGTTTCCCAACGATCAAGCGGAGTTACATGATCCC 4923  
 QY 6597 CCATGTTGTGCAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGTTGTGCAAGTAAGT 6656  
 DB 4924 CCATGTTGTGCAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGTTGTGCAAGTAAGT 4983  
 QY 6657 TGGCCCGAGTTTATCAGTATGTTATGCGAGCACTGCATTAATTTCTTACTGTGTCATGC 6716  
 DB 4984 TGGCCCGAGTTTATCAGTATGTTATGCGAGCACTGCATTAATTTCTTACTGTGTCATGC 5043  
 QY 6717 CATCCGTAAAGTGTCTTTCTGTGACTGGTGAGTCAACCAAGTCAATTTCTGAGATAGT 6776  
 DB 5044 CATCCGTAAAGTGTCTTTCTGTGACTGGTGAGTCAACCAAGTCAATTTCTGAGATAGT 5103  
 QY 6777 GTATGCGGCGACGAGTGTCTTTGCGCGCGTCAATACGGGATATACCGGCCACATA 6836  
 DB 5104 GTATGCGGCGACGAGTGTCTTTGCGCGCGTCAATACGGGATATACCGGCCACATA 5163  
 QY 6837 GCAGAACTTTAAAGTGTCTCATTTGGAAGAGTTCCTGCGGCGGAAACTCTCAAGA 6896  
 DB 5164 GCAGAACTTTAAAGTGTCTCATTTGGAAGAGTTCCTGCGGCGGAAACTCTCAAGA 5223  
 QY 6897 TCTTACCGCTGTGAGATCCAGTTCGATGTAAACCACTCGTGCACCCCACTCATCTTCAG 6956  
 DB 5224 TCTTACCGCTGTGAGATCCAGTTCGATGTAAACCACTCGTGCACCCCACTCATCTTCAG 5283  
 QY 6957 CATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCGCAA 7016  
 DB 5284 CATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCGCAA 5343  
 QY 7017 AAAAGGAATAGGGCGACACGGAATGTTGAATCTCATCTCTCTCTCTCTCTCTCTCAATAT 7076  
 DB 5344 AAAAGGAATAGGGCGACACGGAATGTTGAATCTCATCTCTCTCTCTCTCTCTCTCAATAT 5403  
 QY 7077 ATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGATTTAGA 7136  
 DB 5404 ATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGATTTAGA 5463  
 QY 7137 AAAATAAACAATAGGGGTTCCGCGCAATTTCCCGGAAAGTGCCACCTGACGCTTAAG 7196  
 DB 5464 AAAATAAACAATAGGGGTTCCGCGCAATTTCCCGGAAAGTGCCACCTGACGCTTAAG 5523  
 QY 7197 AAACCAATTAATCATGACATTAACCTATAAAATAGGCGTATCACAGGCGCTTTTCGTC 7256  
 DB 5524 AAACCAATTAATCATGACATTAACCTATAAAATAGGCGTATCACAGGCGCTTTTCGTC 5583  
 QY 7257 TCGCGGTTTCGGTGTATGAGCGTGAACCTCTGACATGAGCTCCCGGAGACGGTCA 7316  
 DB 5584 TCGCGGTTTCGGTGTATGAGCGTGAACCTCTGACATGAGCTCCCGGAGACGGTCA 5643  
 QY 7317 CAGCTTGTCTGTAAGCGGATGCGGAGCAGACAAGCCCGTCAAGGCGGCTCAGCGGGTG 7376  
 DB 5644 CAGCTTGTCTGTAAGCGGATGCGGAGCAGACAAGCCCGTCAAGGCGGCTCAGCGGGTG 5703  
 QY 7377 TTGGCGGGTGTGCGGGCTGGCTTAATGATGCGGATCAAGAGATGTTGATCTGAGAGTGC 7436  
 DB 5704 TTGGCGGGTGTGCGGGCTGGCTTAATGATGCGGATCAAGAGATGTTGATCTGAGAGTGC 5763

QY 7437 ACCATATGCGGTGTGAATATACCGCACAGATGCGTAAGGAGAAAAATACCGCATCAGCGGC 7496  
 DB 5764 ACCATATGCGGTGTGAATATACCGCACAGATGCGTAAGGAGAAAAATACCGCATCAGCGGC 5823  
 QY 7497 ATTGCGCAATTCAGGCTGCGCAACTGTTGGAAAGGCGATCGGTGCGGGCTCTTCGCTAT 7556  
 DB 5824 ATTGCGCAATTCAGGCTGCGCAACTGTTGGAAAGGCGATCGGTGCGGGCTCTTCGCTAT 5883  
 QY 7557 TAGCCAGCTGCGGAAGGGGATGCTGCAAGGCGATTAAGTTGGGTAAAGCCAGGCT 7616  
 DB 5884 TAGCCAGCTGCGGAAGGGGATGCTGCAAGGCGATTAAGTTGGGTAAAGCCAGGCT 5943  
 QY 7617 TTTCCAGTCACGACGTTGTGTAATAACGACGCGCCAGTGCC 7654  
 DB 5944 TTTCCAGTCACGACGTTGTGTAATAACGACGCGCCAGTGCC 5981

RESULT 9  
 US-10-789-938B-1  
 ; Sequence 1, Application US/10789938B  
 ; Publication No. US20050009180A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Yang, Lili  
 ; APPLICANT: Van Parijs, Luk  
 ; APPLICANT: Baltimore, David  
 ; TITLE OF INVENTION: METHOD FOR THE GENERATION OF  
 ; TITLE OF INVENTION: ANTIGEN-SPECIFIC LYMPHOCYTES  
 ; FILE REFERENCE: CALTE.008CP1  
 ; CURRENT APPLICATION NUMBER: US/10/789,938B  
 ; PRIOR APPLICATION NUMBER: 10/317,078  
 ; PRIOR FILING DATE: 2002-12-10  
 ; PRIOR APPLICATION NUMBER: 60/394,803  
 ; PRIOR FILING DATE: 2002-07-08  
 ; PRIOR APPLICATION NUMBER: 60/339,375  
 ; PRIOR FILING DATE: 2001-12-10  
 ; NUMBER OF SEQ ID NOS: 3  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1  
 ; LENGTH: 6254  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: This represents a retroviral vector derived from  
 ; OTHER INFORMATION: the murine stem cell virus.  
 US-10-789-938B-1

Query Match 41.2%; Score 3151.6; DB 19; Length 6254;  
 Best Local Similarity 71.3%; Pred. No. 0;  
 Matches 5077; Conservative 0; Mismatches 889; Indels 1152; Gaps 24;

QY 552 TGAAGAGCCCACTGTAGTTGGCAAGCTAGCTTAAGTAAGCCATTTTCAAGGCAT 611  
 DB 1 TGAAGAGCCCACTGTAGTTGGCAAGCTAGCTTAAGTAAGCCATTTTCAAGGCAT 60  
 QY 612 GGAATAATACATAACTGAGAAATAGAGAGTTTCAAGGTTAGGAACAGAGACAGC- 671  
 DB 61 GGAATAATACATAACTGAGAAATAGAGAGTTTCAAGGTTAGGAACAGAGACAGC- 119  
 QY 672 GAATATGGGCCAAACAGGATATCTGTGGTAAGCAGTTCTCTGCCCGCTCAGGCCCAAGAA 731  
 DB 120 ----- 119  
 QY 732 CAGTTGGAACAGGAGAAATATGGCCAAACAGGATATCTGTGTAAAGCAGTTCTCTGCCCG 791  
 DB 120 -----AGAAATATGGGCCAAACAGGATATCTGTGTAAAGCAGTTCTCTGCCCG 166  
 QY 792 GCTCAGGCGCAAGAACAGATGCTCCCAAGATGCGGTCCCGCTCAGCAGATTCTTAGAGA 851  
 DB 167 GCTCAGGCGCAAGAACAGATGCTCCCAAGATGCGGTCCCGCTCAGCAGATTCTTAGAGA 226  
 QY 852 ACCATCAGATGTTTCCAGGCTGCCCAAGGACCTG-AAATGACCTGTGCGCTTATTGAA 910

Db 227 ACCATCAGATGTTTCCAGGGTGCCTCCCAAGGACCTGAAATGACCTGTGCCTTAATTGAA 286  
Qy 911 CTAACCAATAGTTGGCTTCTCGTTCTGTGTTGCGCGCTTCTGTCTCCCGAGCTCAATAA 970  
Db 287 CTAACCAATAGTTGGCTTCTCGTTCTGTGTTGCGCGCTTCTGTCTCCCGAGCTCAATAA 346  
Qy 971 AAGAGCCACAAACCCCTCACTCGGCGGCCAGTCTCTCGGATAGACTGGCTCGCCGGGTA 1030  
Db 347 AAGAGCCACAAACCCCTCACTCGGCGGCCAGTCTCTCGGATAGACTGGCTCGCCGGGTA 406  
Qy 1031 CCGGTATTCCCAATAAAGCCTTGTGTTTGGCATCGGAATCGTGGACTCGCTCATCTT 1090  
Db 407 CCGGTATTCCCAATAAAGCCTTGTGTTTGGCATCGGAATCGTGGACTCGCTCATCTT 466  
Qy 1091 GGGAGGGTCTCTCAGATTGATTGACTGCCCACTCGGGGTCTTTCAATTTGGAGTTCC 1150  
Db 467 GGGAGGGTCTCTCAGATTGATTGACTGCCCACTCGGGGTCTTTCAATTTGGAGTTCC 526  
Qy 1151 ACCGAGATTGGAGACCCCTGCCAGGACCAACCGACCCCGCGGGAGGTTAAGCTGG 1210  
Db 527 ACCGAGATTGGAGACCCCTGCCAGGACCAACCGACCCCGCGGGAGGTTAAGCTGG 586  
Qy 1211 CCAGCAACTTATGTGTCTGTC-----CGATTGTCTAGTGTCTATGACTCATTTTA 1262  
Db 587 CCAGGGTCTGTTGTTGTTCTGTCTGTCTGTTGCGGTGTTTGTGCGGCTCTAATGTT 646  
Qy 1263 TCGCCTCGCTCGGTACTAGTTAGTAACTAGCTCTGTATCTGGCGGACCCGCTGGTGA 1322  
Db 647 TCGCCTCGCTCTGTACTAGTTAGTAACTAGCTCTGTATCTGGCGGACCCGCTGGTGA 706  
Qy 1323 CTGACGAGTTCCGGAACACCCGCGCAACCTCTGGAGAGCTGCCAGGACTTCGGGGGCC 1382  
Db 707 CTGACGAGTTCTGAACACCCGCGCAACCTCTGGAGAGCTGCCAGGACTTCGGGGGCC 766  
Qy 1383 GTTTTGTGGCGGACCTGAGTCTTAATCCCATCGTTTAGGACTCTTTGGTGCACCC 1442  
Db 767 GTTTTGTGGCGGACCTGAGGAAGGGAGTCTGATGGAATCCGACCCCGTCA----- 819  
Qy 1443 CCCTTAGAGGAGGATATGTGGTCTGTGTAGAGACAGAACTTAAACAGTTCGCGCT 1502  
Db 820 -----GGATATGGTCTGTGTAGAGACAGAACTTAAACAGTTCGCGCT 867  
Qy 1503 CCGTCTGAATTTTGTCTTGGTCTGGACCGAAGCCGCGCGCTTGT---CTGC 1559  
Db 868 CCGTCTGAATTTTGTCTTGGTCTGGAAACCGAAGCCGCGCTTGTCTGCTGCAGCGC 927  
Qy 1560 TGCAGCATCGTCTGTGTGTCTCTGTCTGACTGTGTTCTGTATTTGTCTGAATAATG 1619  
Db 928 TGCAGCATCGTCTGTGTGTCTCTGTCTGACTGTGTTCTGTATTTGTCTGAATAATG 987  
Qy 1620 GGCCGGCTAGACTGTTTACCACTCCCTTAAAGTTTGACTTGGTCTACCTGGAAGATGTC 1679  
Db 988 GGCC-----AGACTGTTTACCACTCCCTTAAAGTTTGACTTGGTCTACCTGGAAGATGTC 1041  
Qy 1680 GAGCGGATCGCTCAACACAGTCTGAGTGTCAAGAAAGAGACGTTGGGTTACCTTCTGC 1739  
Db 1042 GAGCGGATCGCTCAACACAGTCTGAGTGTCAAGAAAGAGACGTTGGGTTACCTTCTGC 1101  
Qy 1740 TCTGAGAAATGGCCAACTTTAACTCGTGGATGGCGCGGAGAGGACCTTTAAACGAGAC 1799  
Db 1102 TCTGAGAAATGGCCAACTTTAACTCGTGGATGGCGCGGAGAGGACCTTTAAACGAGAC 1161  
Qy 1800 CTCATCACCAGGTTAAGATCAAGTCTTTTACCTGGCCCGCATGGACACCCAGACCGAG 1859  
Db 1162 CTCATCACCAGGTTAAGATCAAGTCTTTTACCTGGCCCGCATGGACACCCAGACCGAG 1221  
Qy 1860 GTCCCTCATACGTGACCTGGGAAGCTTGGCTTTTGAACCCCTTCCCTGGGTCAAGGCC 1919  
Db 1222 GTCCCTCATACGTGACCTGGGAAGCTTGGCTTTTGAACCCCTTCCCTGGGTCAAGGCC 1281  
Qy 1920 TTTGTACACCTAAGCCTCCGCTCTCTCTCCATCCGCGCGGTCTCTCCCTTGA 1979  
Db 1282 TTTGTACACCTAAGCCTCCGCTCTCTCTCCATCCGCGCGGTCTCTCCCTTGA 1341

Qy 1980 CTTCTCTGTTGACCCCGCTCGATCTCTCTTTATCCAGCCCTCACTCTCTTCTTAGGC 2039  
Db 1342 CTTCTCTGTTGACCCCGCTCGATCTCTCTTTATCCAGCCCTCACTCTTCTTAGGC 1401  
Qy 2040 GCCCCATATGGCCATATGATCTTATATGGGGCACCCCGCCCTTGTAACTTCCCT 2099  
Db 1402 GCCGAGAT-----CT 1411  
Qy 2100 GACCTCAGATGACAGAGTTTACTTAACAGCCCTCTCTCAAGCTCACTTACAGGCTTCT 2159  
Db 1412 CTGAGAGCGTTAAGCGAGTTTAAACGCGCGCGCGCAAGCT----- 1456  
Qy 2160 ACTTAGTCCAGCAGAAAGTCTGGAGACCTCTGGCGGAGCTTCAAGAAACAACTGGACC 2219  
Db 1457 -----TGACGAATTCGCGCTCTCTCTCCCGCCCTTAACTTACTGGCGAA 1506  
Qy 2220 GACCGGTGTTACCTCACCTTACCAGTTCGGCGACACAGTGTGGTCCGCGACACAGA 2279  
Db 1507 GCGCTTGGAAATAGGCGGTTGCGTGTGTCTATATGTTATTTTCCACCATATGCGCT 1566  
Qy 2280 CTAAGAACTTAGAACCTCGCTGGAAAGGACCTTACACAGTCTCTGCTGACCAACCCACCG 2339  
Db 1567 CTTTGGCAATGTGAGGCGCGGAAACCTGGCCCTGTCTTTCACGAGCATTCCTAGGG 1626  
Qy 2340 CCCTCAAAGTAGACGGCATCGCAGCTTGGATACAGCGCGCCACGTAAGGCTGCGGACC 2399  
Db 1627 GTCTTTCCCTCTCGCAAGGAATGCAAGTCTGTGTAATGTCTGTAAGGAAGCAGTTC 1686  
Qy 2400 CCGGGGTGGACCATCTCTAGACTGCGCATGGGATGGAGCTGTATCATCTCTTCTTGGT 2459  
Db 1687 CTCTG---GAAGCTTCTTGAAGCAAAACAGCTGTGTAGCGACCTTTGCG----- 1733  
Qy 2460 AGCAACAGCTACAGGTGTCCACTCCGACATCCAGCTGACCCAGAGCCCAAGCAGCTGAG 2519  
Db 1734 -----AGCAGCGGAAACCCCCACCTGCGGACAGTGTCTGCGGCCAAAGCAGCTG 1788  
Qy 2520 CGCCAGCGTGGTGAAGAGTGACCATCACTGTGAAGCCAGTCAAGATGTTGGTACTTC 2579  
Db 1789 TATAAGATACACCTGCAAGGCGGCAACACCCAGTGCACGTTGTGAGTTGGATAGTTG 1848  
Qy 2580 TGTAGCTGTTACAGCAGAACGAGTAAAGGCTCCAAAGCTGTGATCTACTGACATC 2639  
Db 1849 TGAAGAGTCA----- 1860  
Qy 2640 CACCGGACACACTGTGTGTGCCAAGCAGATTCAAGGCTAGCGGTAGCGGTACCGACTTCA 2699  
Db 1861 ----- 1860  
Qy 2700 CTTCAACATCAGCAGCCTCCAGCCAGAGGACATCGCCACTTACTTCCAGCAATATAG 2759  
Db 1861 -----AATGGCT 1867  
Qy 2760 CTTCTATCGTCTGTTGCGCCAGGAGCAAGGTGGAATCAACAGAGTGTGCTCAGGATC 2819  
Db 1868 CTCTCAAGCGTATTTCAACAGGGGCTGAAGGATG-----CCAGAAAGTACCCCAAT 1920  
Qy 2820 GGGTGGATCCGGCTCTGTGCTCAGGATCGGAGTCCAACTGTGTGAGAGCGGTGGAGG 2879  
Db 1921 GTATGGATCTGATCTGGGCTC-----GGTGCAATGCTTTACATGTGTTAGT 1971  
Qy 2880 TGTGTGCAACCTTGGCGGCTCTCTGCGCTGTCTGTCTCGCATCTGCTTTCGATTTCA 2939  
Db 1972 CGAGGTTAAACAGCTCTAGC---CCCCGAAACACGGGACGCTGTTTCTTCTTGA 2028  
Qy 2940 CACATATTGGATGTTGGGTGAGACAGGCACTCGGAACAGGTCTTGTAGTGGATGGAGA 2999  
Db 2029 AACACGATGATAATGGCCACAACCAAGGCGGAGGAGCTGTTCACCGGGGTGTGCCCCA 2088  
Qy 3000 AATTTCATCCAGATAGCAGTACGATTAATATGCGCGCTCTCTAAAGGATAGATTTACAT 3059  
Db 2089 TCCTGGTCCAGCTGG----- 2103

QY 3060 ATCGCGAGACAAAGCGCAAGAACACATTTGTTCTCTGCAAAATGGACAGCTTGAGACCGCAAGA 3119  
 Db 2104 -----ACGGCGAGCTGAACGSCCAACAAGTTTCAGGGTTCGGGCGAGG 2146  
 QY 3120 CACCGGGGTCTATTTTGTGCAAGCCCTTACTTTGGCTTCCCTCGTGTGTTGCTTATTTGGGG 3179  
 Db 2147 CGAGGGCGATGCGCACCTACGGAAGCTGACCCCTGAAGTTTCATCTGCAACACCGGCAAGCT 2206  
 QY 3180 CCAAGGGACCCCGGTACCGTCTCCAGTCTTAAGCCACCAAGCGCGACCGCGGACC 3239  
 Db 2207 GCCGTGCTTGGCCACCCCT-----CGTGACC 2234  
 QY 3240 ACCAAACACCGGGCCCAACATCGCTCGAGCCCTCTCCCTCGGCGCCACAGCGGCTCG 3299  
 Db 2235 ACCCTGACCTACGGGTGAGTGTCT--CAGCGCTACCCCGACACATGAAGACAGCAGC 2292  
 QY 3300 GCGAGCGGGGGCGCAGTGACACAGAGGGGTGGACTTGGCCCTGGATCCCAAACT 3359  
 Db 2293 AC-----TTCTTCAAGTCGGCCATGCCCGAAGG 2320  
 QY 3360 CTGCTACCTGCTGGATGGAATCCTCTTCACTATGCTGTCTATCTCACTGCTGCTTGTCT 3419  
 Db 2321 CTACGTCGAGGACGACCATCTTCTTCAAGGACGACGCAACTACAAGACCCGCGCCA 2380  
 QY 3420 GAGAGTGAAGTTTCAGCAGAGCGCAGAGCCCGCGTACACAGCAGGGCCAGAACAGCT 3479  
 Db 2381 GGTGAAGTTGAGGGCGACACCTGTTGAACCGCATCGAGCTGAAGGGCATCGACTTCAA 2440  
 QY 3480 CTATAACGAGCTCAATCTTAGGACGAGAGGAGTACGATGTTTGGACAAGAGAGCTGG 3539  
 Db 2441 GGAGGACGGCAACATCTCGGGCACAAGCTGAGTACAACCTACAACAGCCACAACGCTTA 2500  
 QY 3540 CCGGACCTGAGATGGGGGAAAGCCGAGAGGAGAACCTCAGGAAGGCTGTACAA 3599  
 Db 2501 TATCATGGCCGCAAGACGAAGAAGCGCATCAAGCGCACTTCAAGATCCGCCACAACAT 2560  
 QY 3600 TGAACCTGCAGAAAGATAAGATGCGGAGGCTTACAGTGAGATTGGGATGAAGGCGAGCG 3659  
 Db 2561 CAGGAGCGGCTGACCTGCGGACCACTACCGAGAGAACACCCCATCGGCGAGG 2620  
 QY 3660 CCGGAGGGCAAGGGGCAAGATGGCTTTTACCAGGCTTCAGTACAGCCACCAAGGACAC 3719  
 Db 2621 CCGCGTGTCTGCTGCCGCAACCACTACCTGAGCACCAGTCCGCTCAGCAAGACCC 2680  
 QY 3720 CTACGAGCCCTTACATCGAGCCCTGCCCCCTCGTAACTCGAGCGGCGCGGATCC 3779  
 Db 2681 CAAACGAGAACCGGATCAATGGTCTCTGGAGTTCTGTGACCGCGCGCGGATCACTCA 2740  
 QY 3780 GGATTAGTCCAAATTTGTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTGTACTCA 3839  
 Db 2741 CGGCATGAGACGAGTGTACAGTAAGTCACTGCGAGCCAGCTT----- 2785  
 QY 3840 CAATATCACAGCTGAAGCTTATAGAGTACGAGCCATAGATAAAATAAAGATTTTATTT 3899  
 Db 2786 -----ATCGATAAAATAAAGATTTTATTT 2810  
 QY 3900 AGTCTCCAGAAAAGGGGGGAATGAAGAAGCCCACTGTAGTTTGGGCAAGCTAGCTTAA 3959  
 Db 2811 AGTCTCCAGAAAAGGGGGGAATGAAGAAGCCCACTGTAGTTTGGGCAAGCTAGCTTAA 2870  
 QY 3960 GTAACGCCATTTTGAAGCATGGAATAACATACTGAGAAATAGAGAAGTTTCAGATCAA 4019  
 Db 2871 GTAAAGCCATTTTGAAGCATGGAATAACATACTGAGAAATAGAGAAGTTTCAGATCAA 2930  
 QY 4020 GGTTAGGAACAGAGACAGCAGAAATATGGGGCAACAGGATATCTGTGTGAAGCAGTTTC 4079  
 Db 2931 GGTTAGGAACAGAGACAGC----- 2951  
 QY 4080 CTGCCCCCTCAGGGCCAAAGAACAGTTGGAAACAGGAGATATGGGCAACAGGATATCT 4139  
 Db 2952 -----AGAAATATGGGCAACAGGATATCT 2976  
 QY 4140 GTGGTAAGCAGTTCTGCCCCGCTCAGGGCCAAAGAACAGATGGTCCCCCAGATGCGGTCC 4199

Db 2977 GTGGTAAGCAGTTCTGCCCCGCTCAGGGCCAAAGAACAGATGGTCCCCCAGATGCGGTCC 3036  
 QY 4200 CGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGGCCCAAGACCTG-AA 4258  
 Db 3037 CGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGGCCCAAGACCTGAAA 3096  
 QY 4259 ATGACCCCTGTGCTTATTTGAACCAATCAGTTCGCTTCTGCTTCTGTTGCGCGC 4318  
 Db 3097 ATGACCCCTGTGCTTATTTGAACCAATCAGTTCGCTTCTGCTTCTGTTGCGCGC 3156  
 QY 4319 TTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGCGCGCCAGTCTCTCC 4378  
 Db 3157 TTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGCGCGCCAGTCTCTCC 3216  
 QY 4379 GATAGACTGCGTGGCCGCGGTACCGTGTCTCAATAAACCCCTCTTGCAAGTTGCATCCA 4438  
 Db 3217 GATAGACTGCGTGGCCGCGGTACCGTGTCTCAATAAACCCCTCTTGCAAGTTGCATCCA 3276  
 QY 4439 CTGCTGCTCGCTCTCTTGGGAGGGTCT--CTCTGAGTGATTGACTACCCGTGAGCGG 4497  
 Db 3277 CTGCTGCTCGCTCTCTTGGGAGGGTCTCTCTGAGTGATTGACTACCCGTGAGCGG 3336  
 QY 4498 GGTCTCTTTCAGTTTCTCCACCTACACAGGTCTCTCAATAATCTCTGATGTGCGCGAGGA 4557  
 Db 3337 GGGTCTT----- 3343  
 QY 4558 CTCCTCAGCCCGGTTTTTGTATTAATAAATGCAAGAACAGTGTTCCTTCAAGCCAG 4617  
 Db 3344 ----- 3343  
 QY 4618 ACTACATCTGACTCTCGCTTTTATAAAGAAATGTTGAAGGCTCTGTGGACTATCTGCC 4677  
 Db 3344 ----- 3343  
 QY 4678 ACACGACTTTTAAAGATTTTATGCTCTCTGATGAGGAAATTTAGTCAATCTATCTCGT 4737  
 Db 3344 ----- 3343  
 QY 4738 CTATTTTGTGCTTCTCGTATTTTAAATTTCTAGTTTGCATCCCTTCTGAGAGCAC 4797  
 Db 3344 ----- 3343  
 QY 4798 GCGGATTGCAGAGTAGTTAATCTCTGAGGGCAGGCTCTGTGAAAAAGTTGCTTGGGCT 4857  
 Db 3344 ----- 3343  
 QY 4858 CAGTGTGAGATTTTCCCATAAAAAGGGTCTCTGCCCCCTGTGTACAGACAGATCGGAATCT 4917  
 Db 3344 ----- 3343  
 QY 4918 AGAGTGCATATCTCAGAGTCCCCCGGGTCCGGGGCTCTGATCTCAGGGGATCTTTGCTTA 4977  
 Db 3344 ----- 3343  
 QY 4978 GAGATCCTCTAGCCCGAGCATCGTGGCGGGTACCGAGCTCGAATTCGTAATCATGCT 5037  
 Db 3344 -----TCAGTATTCGTAATCATGCT 3363  
 QY 5038 CATAGCTGTTTCTGTGAAATTTGTTATCCGCTCACAATTCACACAACATACGAGCGG 5097  
 Db 3364 CATAGCTGTTTCTGTGAAATTTGTTATCCGCTCACAATTCACACAACATACGAGCGG 3423  
 QY 5098 GAAGCATAAAGTGTAAAGCTCGGGGTGCTTAATGATGAGTGAAGTCACTCAATTAATTTGCT 5157  
 Db 3424 GAAGCATAAAGTGTAAAGCTCGGGGTGCTTAATGATGAGTGAAGTCACTCAATTAATTTGCT 3483  
 QY 5158 TCGCTCTACTGCGCTTTCAGTCCGGAAACCTGTCGTCAGCTGCAATTAATTAATCG 5217  
 Db 3484 TCGCTCTACTGCGCTTTCAGTCCGGAAACCTGTCGTCAGCTGCAATTAATTAATCG 3543  
 QY 5218 GCCAACGCGCGGAGAGCGGTTTTCGTTATTTGGCGCTCTTTCGCTTCTCTCGCTCACTG 5277

Db 3544 GCCAACGGCGGGGAGAGGCGGTTTGGCTATTGGGCGCTCTTCCGCTCTCTCGCTCACTG 3603  
Qy 5278 ACTCGCTCGCTCGGTCGTTGCGGTGCGGCGAGCGGTATCAGCTCACTCAAAGCGGTAA 5337  
Db 3604 ACTCGCTCGGCTCGGTCGTTGCGGTGCGGCGAGCGGTATCAGCTCACTCAAAGCGGTAA 3663  
Qy 5338 TACGGTTATCCACAGAAATCAGGGATTAACGAGAAAGAAACATGTGAGCAAAAAGCCAGC 5397  
Db 3664 TACGGTTATCCACAGAAATCAGGGATTAACGAGAAAGAAACATGTGAGCAAAAAGCCAGC 3723  
Qy 5398 AAAAGGCGAGAAACCGTAAAGGCGCGGTTGCTGGGGTTTTTCCATAGGCTCCGCCCCC 5457  
Db 3724 AAAAGGCGAGAAACCGTAAAGGCGCGGTTGCTGGGGTTTTTCCATAGGCTCCGCCCCC 3783  
Qy 5458 CTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGCGAAACCCGACAGGACTAT 5517  
Db 3784 CTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGCGAAACCCGACAGGACTAT 3843  
Qy 5518 AAAGATACGAGCGGTTTCCCGCTGGAAGCTCCCTCGTGGCGCTCTCTGTTCCGACCCCTGC 5577  
Db 3844 AAAGATACGAGCGGTTTCCCGCTGGAAGCTCCCTCGTGGCGCTCTCTGTTCCGACCCCTGC 3903  
Qy 5578 CGCTTACCGGATACCTGTCGCGCTTCTCCCTCGGGAAGCGTGGCGCTTCTCATAGCT 5637  
Db 3904 CGCTTACCGGATACCTGTCGCGCTTCTCCCTCGGGAAGCGTGGCGCTTCTCATAGCT 3963  
Qy 5638 CACGCTGTAGGTATCTCAGTTTCGGGTGAGTTCGTTCCGTCGAAGCTGGGCTGTGCAACG 5697  
Db 3964 CACGCTGTAGGTATCTCAGTTTCGGGTGAGTTCGTTCCGTCGAAGCTGGGCTGTGCAACG 4023  
Qy 5698 AACCCCGGTTTACGCCGACCGCTGCGCTTTATCCGGTAACATCGTCTTGAGTCCAAAC 5757  
Db 4024 AACCCCGGTTTACGCCGACCGCTGCGCTTTATCCGGTAACATCGTCTTGAGTCCAAAC 4083  
Qy 5758 CGGTAAACGACGACTTACGCACTGCGCAGCAGCACTGGTAAACAGNTAGCAGGCGA 5817  
Db 4084 CGGTAAACGACGACTTATCGCCACTGCGCAGCAGCACTGGTAAACAGNTAGCAGGCGA 4143  
Qy 5818 GGATGTAGGCGGTGTACAGAGTTCTTGAAGTGGTGGCCCTAACTACGGCTACACTAGAA 5877  
Db 4144 GGATGTAGGCGGTGTACAGAGTTCTTGAAGTGGTGGCCCTAACTACGGCTACACTAGAA 4203  
Qy 5878 GGA CAGTATTGGTATCTGGCTCTGTGAAGCCAGTTACCTTCGGA AAAAGAGTTGGTA 5937  
Db 4204 GGA CAGTATTGGTATCTGGCTCTGTGAAGCCAGTTACCTTCGGA AAAAGAGTTGGTA 4263  
Qy 5938 GCTCTTGATCCGGCAAAACACCGCTGGTAGCGGTGGTTTTTTTGGTTGCAAGCAGC 5997  
Db 4264 GCTCTTGATCCGGCAAAACACCGCTGGTAGCGGTGGTTTTTTTGGTTGCAAGCAGC 4323  
Qy 5998 AGATTACGCGAGAAAAAGGATCTCAAGAGATCTTTTGATCTTTTCTACGGGGTCTG 6057  
Db 4324 AGATTACGCGAGAAAAAGGATCTCAAGAGATCTTTTGATCTTTTCTACGGGGTCTG 4383  
Qy 6058 ACGCTCAGTGAACGAAAACTCAGGTTAAGGGAATTTTGGTCATGAGATTAACA AAAAGGA 6117  
Db 4384 ACGCTCAGTGAACGAAAACTCAGGTTAAGGGAATTTTGGTCATGAGATTAACA AAAAGGA 4443  
Qy 6118 TCTTACCTAGATCCTTTTAAATTA AAAA TGAAGTTTTTAAATCAATCTAAAGTATATATG 6177  
Db 4444 TCTTACCTAGATCCTTTTAAATTA AAAA TGAAGTTTTTAAATCAATCTAAAGTATATATG 4503  
Qy 6178 AGTAAATCTGCTGACAGTTTACCATGCTTAATCAGTAGGACCTATCTCCGAGATCT 6237  
Db 4504 AGTAAATCTGCTGACAGTTTACCATGCTTAATCAGTAGGACCTATCTCCGAGATCT 4563  
Qy 6238 GTCTATTCTGTTCACTAGTTGCTGACTCCCGCTCGTGTAGATAACTACGATACGGG 6297  
Db 4564 GTCTATTCTGTTCACTAGTTGCTGACTCCCGCTCGTGTAGATAACTACGATACGGG 4623  
Qy 6298 AGGCTTACCATCTGGCCCAAGTGTGCAATGATACCGCGAGAACCCAGCTCACCGGCTC 6357  
Db 4624 AGGCTTACCATCTGGCCCAAGTGTGCAATGATACCGCGAGAACCCAGCTCACCGGCTC 4683

Qy 6358 CAGATTTATCAGCAATAAAAACAGCCAGCCGGAAGCGGCGAGCGAGAGTGGTCTCTGCAA 6417  
Db 4684 CAGATTTATCAGCAATAAAAACAGCCAGCCGGAAGCGGCGAGCGAGAGTGGTCTCTGCAA 4743  
Qy 6418 CTTTATCCGCTCCATCCAGTCTATTAAATTTGTTGCGGGAAGCTAGAGTAAAGTATTCGC 6477  
Db 4744 CTTTATCCGCTCCATCCAGTCTATTAAATTTGTTGCGGGAAGCTAGAGTAAAGTATTCGC 4803  
Qy 6478 CAGTTAATAGTTTTCGCAACAGTTGTTGCCATTGCTACAGGC - TCGTGGTGTCAAGCTCGT 6536  
Db 4804 CAGTTAATAGTTTTCGCAACAGTTGTTGCCATTGCTACAGGCATGTTGGTGTCAAGCTCGT 4863  
Qy 6537 CGTTTGGTATGCGTTTCAATCAGCTCCGTTTCCCAACAGTCAAGCGAGTTCATCATGATCCC 6596  
Db 4864 CGTTTGGTATGCGTTTCAATCAGCTCCGTTTCCCAACAGTCAAGCGAGTTCATCATGATCCC 4923  
Qy 6597 CCATGTTTGTGCAAAAAAGCGGTTAGCTCTTCGCTCCCTCCGATCGTTGTCAAGATTAAGT 6656  
Db 4924 CCATGTTTGTGCAAAAAAGCGGTTAGCTCTTCGCTCCCTCCGATCGTTGTCAAGATTAAGT 4983  
Qy 6657 TGGCCGAGGTTTATCACTCATGTTATGGCAGCACTGCATAATCTCTTACTGTCAATGC 6716  
Db 4984 TGGCCGAGGTTTATCACTCATGTTATGGCAGCACTGCATAATCTCTTACTGTCAATGC 5043  
Qy 6717 CATCCGTAAGATGCTTTTCTGTGATCGTGTGAGTACTCAACCAAGTCAATCTTGAGAATAAGT 6776  
Db 5044 CATCCGTAAGATGCTTTTCTGTGATCGTGTGAGTACTCAACCAAGTCAATCTTGAGAATAAGT 5103  
Qy 6777 GTATGCGGCGACCGAGTTGCTCTTCCCGGCGTCAATAACGGGATTAATACCGGCGCACATA 6836  
Db 5104 GTATGCGGCGACCGAGTTGCTCTTCCCGGCGTCAATAACGGGATTAATACCGGCGCACATA 5163  
Qy 6837 GCAGAACTTTAAAGTCTCATCATCTGAAAAAGCTTCTTCGGGGCGAAAACTCTCAAGGA 6896  
Db 5164 GCAGAACTTTAAAGTCTCATCATCTGAAAAAGCTTCTTCGGGGCGAAAACTCTCAAGGA 5223  
Qy 6897 TCTTACCGCTGTTGAGATCCAGTTTCGATGTAAACCACTCGTGCAACCACTGATCTTTCAG 6956  
Db 5224 TCTTACCGCTGTTGAGATCCAGTTTCGATGTAAACCACTCGTGCAACCACTGATCTTTCAG 5283  
Qy 6957 CATCTTTTACTTTTACACAGCGTTTCTGGGTGAGCAAAAAACGAAAGGCAAAATCCGCGAA 7016  
Db 5284 CATCTTTTACTTTTACACAGCGTTTCTGGGTGAGCAAAAAACGAAAGGCAAAATCCGCGAA 5343  
Qy 7017 AAAAGGGAATAAGGGCGACACGGAATGTTGAATCTCATCTCTCTTTTCAATATT 7076  
Db 5344 AAAAGGGAATAAGGGCGACACGGAATGTTGAATCTCATCTCTCTTTTCAATATT 5403  
Qy 7077 ATTGAAGCATTTTATCAGGTTTATTTGCTCATGAGCGGATACATATTTGAATGTATTTAGA 7136  
Db 5404 ATTGAAGCATTTTATCAGGTTTATTTGCTCATGAGCGGATACATATTTGAATGTATTTAGA 5463  
Qy 7137 AAAATAAATAAATAAGGGTTCCGCGCA CATTTTCCCGAAAAAGTGCCACCTGACGCTTAAG 7196  
Db 5464 AAAATAAATAAATAAGGGTTCCGCGCA CATTTTCCCGAAAAAGTGCCACCTGACGCTTAAG 5523  
Qy 7197 AAAACCATTTATCATGACATTAACCTTATAAATAGCGTATCCAGGCGCTTTTCGTC 7256  
Db 5524 AAAACCATTTATCATGACATTAACCTTATAAATAGCGTATCCAGGCGCTTTTCGTC 5583  
Qy 7257 TCCGCGGTTTTCGTTGATGAGCGGTGAAAACTCTGACACATGACGCTCCCGAGACGGTCA 7316  
Db 5584 TCCGCGGTTTTCGTTGATGAGCGGTGAAAACTCTGACACATGACGCTCCCGAGACGGTCA 5643  
Qy 7317 CAGCTTGTCTGTAAGCGGATGCGGGAGCAGACAAGCCCGTTCAGGGCGCGTTCAGCGGTG 7376  
Db 5644 CAGCTTGTCTGTAAGCGGATGCGGGAGCAGACAAGCCCGTTCAGGGCGCGTTCAGCGGTG 5703  
Qy 7377 TTAGCGGGGTTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTTGTA CTGAGAGTGC 7436  
Db 5704 TTAGCGGGGTTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTTGTA CTGAGAGTGC 5763

QY 7437 ACCATATGCGGTGTGAATAACCGCACAGATGCGTAAAGAGAGAAATACCGCATCAGCGCC 7496  
 Db 5764 ACCATATGCGGTGTGAATAACCGCACAGATGCGTAAAGAGAGAAATACCGCATCAGCGCC 5823  
 QY 7497 ATTGCGCATTCAGGCTGCGCAACTGTGTGGGAAGGCGATCGGTGCGGGCTCTTTCGCTAT 7556  
 Db 5824 ATTGCGCATTCAGGCTGCGCAACTGTGTGGGAAGGCGATCGGTGCGGGCTCTTTCGCTAT 5883  
 QY 7557 TAGCGCAGCTGCGGAAGGGGATGTCTGCAAGGCGATTAAGTTGGGTAACGCCAGGCT 7616  
 Db 5884 TAGCGCAGCTGCGGAAGGGGATGTCTGCAAGGCGATTAAGTTGGGTAACGCCAGGCT 5943  
 QY 7617 TTTTCCAGTCAGCAGCTTTGTAATAACGACGCCAGTGCC 7654  
 Db 5944 TTTTCCAGTCAGCAGCTTTGTAATAACGACGCCAGTGCC 5981

RESULT 10  
 US-09-808-743-2  
 ; Sequence 2, Application US/09808743  
 ; Patent No. US2002006871A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: PEDERSEN, Peter  
 ; APPLICANT: MATHUPALA, Saroj  
 ; TITLE OF INVENTION: ARREST OF PROLIFERATION OF HIGHLY GLYCOLYTIC TUMORS  
 ; FILE REFERENCE: JHUI720-1  
 ; CURRENT APPLICATION NUMBER: US/09/808,743  
 ; CURRENT FILING DATE: 2001-03-14  
 ; PRIOR APPLICATION NUMBER: US 60/189,222  
 ; PRIOR FILING DATE: 2000-03-14  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 2  
 ; LENGTH: 6444  
 ; TYPE: DNA  
 ; ORGANISM: Artificial sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Cloning vector pLXRN  
 US-09-808-743-2

Query Match 40.5%; Score 3103.4; DB 9; Length 6444;  
 Best Local Similarity 71.0%; Pred. No. 0;  
 Matches 4779; Conservative 0; Mismatches 1466; Indels 486; Gaps 31;

QY 549 TTTTGAAGACCCACCTGTAGGTTTGGCAAGTAGCTTAAGTAAGCCCATTTTGAAGG 608  
 Db 174 TTTTGAAGACCCACCCGTAGG--TGGCAAGTAGCTTAAGTAAGCCCATTTTGAAGG 231  
 QY 609 CATGG-AAAATACATACTGAGATAGAGATTCAGATCAAGGTTAGGAAACAGAGAGAC 667  
 Db 232 CATGGAAAATACATACTGAGATAGAGATTCAGATCAAGGTTAGGAAACAGAGAAAC 291  
 QY 668 AGCAGAAATAGGCCCAACAGGATATCTGTGTAGCAGTCTCTGCCCC-GCTCAGGGCC 726  
 Db 292 AGCTGAATA-CCAAACAGGATATCTGTGTAGCAGTCTCTGCCCCGTGAGGCC 347  
 QY 727 AAGAACAGTTGGAACAGGAGAAAT-ATGGGCGCAACAGGATATCTGTGTAGCAGTCTCT 785  
 Db 348 AAGAACAGATGAGACAGCTGAGTGTAGTGGGCGCAACAGGATATCTGTGTAGCAGTCTCT 407  
 QY 786 GCCCGGCTCAGGGCCAGAACAGATAGTCCCGATGCGTCCCGCCCTCAGCAGTTTC 845  
 Db 408 GCCCGGCTCAGGGCCAGAACAGATAGTGGTCCCGATGCGTCCCGCCCTCAGCAGTTTC 467  
 QY 846 TAGAGAACCATCAGATGTTTTCAGGGTGCCCAAGGACCTG-AAATGACCTGTGCTTTA 904  
 Db 468 TAGTGAATCATCAGATGTTTTCAGGGTGCCCAAGGACCTGAAATGACCTGTACCTTA 527  
 QY 905 TTTGAACCAATCAGTTTCGCTTCTCGGTTCTCGGTTCTGTGTCGGGCGCTTCTGCTCCGAGCT 964  
 Db 528 TTTGAACCAATCAGTTTCGCTTCTCGGTTCTGTGTCGGGCGCTTCTGCTCCGAGCT 587

QY 965 CAATAAAGAGAGCCCAACCCCTCACTGGCGCGCAGTCTCTCGATAGACTGCGTCCG 1024  
 Db 588 CAATAAAGAGAGCCCAACCCCTCACTGGCGCGCAGTCTCTCGATAGACTGCGTCCG 647  
 QY 1025 CGGCTACCGGTATTCCTCAATAAAGCTCTTGTGTGTTTGCATCCGAATCGTGTGACTG 1084  
 Db 648 CGGCTACCGGTATTCCTCAATAAAGCTCTTGTGTGTTTGCATCCGAATCGTGTGACTG 707  
 QY 1085 ATCTCTTGGAGGGTCTCTCTCAGATTGATTGATGCTGCCCACTCGGGGTCTTTCAATTGGA 1144  
 Db 708 TTCTCTTGGAGGGTCTCTCTCAGATTGATTGATGCTGCCCACTCGGGGTCTTTCAATTG 767  
 QY 1145 GGTTCACAGGATTGGAGACCCCTGCCAGGAGCACCGACCCCGCGCGGAGGTA 1204  
 Db 768 GGTTCGTCCGGATTGGAGACCCCTGCCAGGAGCACCGA-CCACACCGGAGGTA 826  
 QY 1205 AGCTGGCAGCAACTTATCTGTGTCTGTCGCAATGCTGTAGTGTCTATGACTGATTTATG 1264  
 Db 827 AGCTGGCAGCAACTTATCTGTGTCTGTCGCAATGCTGTAGTGTCTATGATTTATG 886  
 QY 1265 CGCTCTGCGTACGTAGTTAGTAACTGCTCTGATCTGCTGCGGACCCGCTGTGGAAC 1324  
 Db 887 CGCTCTGCGTACGTAGTTAGTAACTGCTCTGATCTGCTGCGGACCCGCTGTGGAAC 946  
 QY 1325 GACGAGTTTCGAACACCCCGCGCAACCTCTGGGAGACGTCCTCCAGGAGACTTCGGGGCCGT 1384  
 Db 947 GACGAGTTTCGAACACCCCGCGCAACCTCTGGGAGACGTCCTCCAGGAGACTTCGGGGCCGT 1006  
 QY 1385 TTTTGTGCGGACCTGAGTCTCTAAATCCCGATCGTTTAGGACTCTTTGTTGTCACCC 1444  
 Db 1007 TTTTGTGCGGACCTGAGTCTCTGAGGAGGAGTCTGATGTGGAATCCGACCC----- 1052  
 QY 1445 CTTAGAGAGGATATGTGTTCTGTTAGGAGACGAGACCTTAAACAGACTTCGCCCTCC 1504  
 Db 1053 -----CGTCAGGATATGTGTTCTGTTAGGAGACGAGACCTTAAACAGACTTCGCCCTCC 1107  
 QY 1505 GTCTGAATTTTGTCTTGGTTGGGACCGAAGCCGCGCGCGCTCTTGT---CTGCTG 1561  
 Db 1108 GTCTGAATTTTGTCTTGGTTGGGACCGAAGCCGCGCGCTCTTGTCTGCTCAGCGCTG 1167  
 QY 1562 CAGCATCGTTGTGTCTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1621  
 Db 1168 CAGCATCGTTGTGTCTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1227  
 QY 1622 CCGGGCTAGACTGTTTACCTCCCTTAACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1681  
 Db 1228 CC-----AGACTGTTTACCTCCCTTAACTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1281  
 QY 1682 GCGGATCGCTCAACACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1741  
 Db 1282 GCGGATCGCTCAACACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1341  
 QY 1742 TGCAGATGGCCAACTTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1801  
 Db 1342 TGCAGATGGCCAACTTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1401  
 QY 1802 CATCACCCAGGTTAGATCAAGTCTTTCACCTGCGCGCATGACACCCAGACAGCT 1861  
 Db 1402 CATCACCCAGGTTAGATCAAGTCTTTCACCTGCGCGCATGACACCCAGACAGCT 1461  
 QY 1862 CCCCTACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1921  
 Db 1462 CCCCTACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1521  
 QY 1922 TGTACACCTTAAGCTCCGCTCTCTTCTCTCCATCCGCGCGCTCTCTCCCTCTTGAACC 1981  
 Db 1522 TGTACACCTTAAGCTCCGCTCTCTTCTCTCCATCCGCGCGCTCTCTCCCTCTTGAACC 1581  
 QY 1982 TCTCTGTTGACCCGCTCTGCT 2041  
 Db 1582 TCTCTGTTGACCCGCTCTGCT 1641  
 QY 2042 CCCCATATGGCCATATGAGATCTTATATGGGGACACCCCGCGCTTGTAAACTTCCCTCTGA 2101



Db 1642 CGGAATTCGTTAACTCGAGGATCCACTAGTAACGGCCGCCAGTGTGCTGGAAATTAATTCG 1701  
Qy C--CCTGACATGACAAGAGTTACTAACAGCCCTCTCTCCAAGCTCACTTACAGGCTTCT 2159  
Db CTGCTCGAGGGCGCGGTGTGGGGTCACTACTCCCTCTCAAAGCGGGCATGACTTCT 1761  
Qy 2160 ACTTAGTCCAGCAGAACTCTGGAGACTCTCTGGCGGCAGCCTTACCAAGAACAACTGGACC 2219  
Db 1762 CGCCTAAGATTGTCAAGTTTCCAAAACGAGGAGGATTTGATATTCACTGCGCCGCGTG 1821  
Qy 2220 GACCGGTGTACTCACCCTTACCGAGTCGGGACACAGTGTGGTCCGCGCACACAG 2279  
Db 1822 ATGCCCTTTGAGGGTGGCGCGCTCCATCTGGTCAAGAAAGACAATCTTTTGTGTCAAGC 1881  
Qy 2280 CTAAGAACCTAGAACCTCGCTGGAAAGACCTTACACAGTCTGCTGACCAACCCCAACCG 2339  
Db 1882 TTGAGGTGGCAGGCTTGAGATCTGGGCCATACACTTGAGTGACAAATGACATCACTTTG 1941  
Qy 2340 CCTCAAAAGTAGCGGCATCGAGCTTGGATACACGCCGCCACAGTGAAGGCTGCCGACC 2399  
Db 1942 CTTTCTCTCACAGGTGTCCACTCCAGGTCCAACTCGAGTCCGATCGATGATCT 2001  
Qy 2400 CCGGGGTGGACCATCTCTAGACTGCGCATGGGATGGAGCTGTATCATCTCTTCTTGGT 2459  
Db 2002 AGGGCGGCCAATTGGCCCTCTCCCTCCCGCCCTTAACGTGTACTGGCCGAAGCGCTT 2061  
Qy 2460 AGCAACAGCTACAGGTGTCCACTCGACATCCAGCTGACCCAGAGCCCAAGCAGCTGAG 2519  
Db 2062 GGAATAGCGCGTGTGTGTGTCTATATGTGATTTTCCACATATTTGCCGTCTTTTGG 2121  
Qy 2520 CGCCAGCGTGGGTGACAGAGTGACCATCACCTGTAAAGGCCAGTCAAGATGTGGGTACTTC 2579  
Db 2122 CAATGTAGGGCCCGGAAACCTGGCCCTGTCTCTTGACGAGCATTCCTAGGGGTCTTTC 2181  
Qy 2580 TGTAGCTTGGTACAGCAGAAAGCCAGGTAAAGCTTCAAAGCTGTGTGATCTACTGACATC 2639  
Db 2182 -----CCCTCTCGCCAAAGGAATGCAAGGTCTGTGTAATGTGTGTAAGGAAGAGTTC 2234  
Qy 2640 CACCGGCACACTGTGTGCCAAGCAGATTACAGCGGTAGCGGTAGCGGTACCGACTTCAC 2699  
Db 2235 CTCTGGAAGCTTCTTGAAGACAACAACAGTCTGTAGCAGCCCTTTGACGAGCGGGAACC 2294  
Qy 2700 CTTCAACATCAGAGCTCTCCAGCCAGAGAGACATCGCCACCTACTACTGCCAGCAATATAG 2759  
Db 2295 CCCCACTGGCGACAGGTGCCTCTCGGCCCAAGCCAGCT-----GTATAGATACA 2347  
Qy 2760 CTTCTATCGTCTGTTGGCCAAAGGACAAGGTGGAATCAAAAGAGTGGCTCAGGATC 2819  
Db 2348 CTTGAAAGGGCGCACAAACCCCACTGCCACGCTTGTGAGTTGGATAGTTGTGGAAGAGTC 2407  
Qy 2820 GGGTGGATCCGGCTCTGTGTGCTCAGGATCGAGGTCCAACTGGTGGAGCGGTGGAG 2879  
Db 2408 AAATGGCTCTCTCAAGGTAGTCAACAAGGGGTGAAGAGTGGCCGAAGGTACCCCAT 2467  
Qy 2880 TGTGTGCAACTGGCCGGTCCCTCGCGCTCTCTGCTCGCATCTGGCTTCGATTTTCAAC 2939  
Db 2468 TGTATGGGAATCTGATCTGGGGCTCGGTGACATGCTTTACATCG-----TTTAG 2519  
Qy 2940 CACATATTGGATAGTTGGGTGAGACAGGCACTTGGAAAGGTCTTGAAGTGGATGGAGA 2999  
Db 2520 TCGAGGTTAAGAAAGCTTAGGGCCCGCCGAACCAACCGGGGAGCTGTGTTTCTTTT---GA 2575  
Qy 3000 AATTTCATCCAGATAGCATGATTAATACTATGCGCGCTCTCTAAGGATAGATTACAT 3059  
Db 2576 AAAACAGATAGTGTGCAACAACCCCGGATAATTCTCGAGCCAAATATGGATCG 2635  
Qy 3060 ATCGGACACAAAGCAGAAACATTTCTCTGCAATGACAGCCCTGACACCGCA--A 3117  
Db 2636 GCCATTGAAAGATGGAATTCAGCGAGTTCTCGGGCGCTTGGTGGAGAGGCTATTC 2695  
Qy 3118 GACACCGGGTCTATTTTGTGCAAGCCTTTACTTCGGCTTCCCTGCTTTTATTTGG 3177

Db 2696 GGCTATGATGGGCACAAACAGACAATCGGCTGCTCTGTATGTCGCGCGTGTTCGGGTGTCA 2755  
Qy 3178 GGCCAAAGGAGCCCGGTCAACGCTCTCCAGTGTCTAAGCCACACAGCAGCGCAGCGCGCA 3237  
Db 2756 GCGCAGGGCGCCCGGTCTTTTGTCAAGACCGACCTGTCCGTTGCC--TGAATGNA 2812  
Qy 3238 CCACAACACACGGCGCCACCATCGGTTCGAGCCCTGTCTCTGCGCCGAGAGCGGT 3297  
Db 2813 CTGAGGACGAGGAGCGCGCTATCGGTCTGGCCACGACGGCGTTCCTTGTGCGAGCT 2872  
Qy 3298 CGCCAGCGGGGGGGCGAGTGCACACAGAGGGGTGGACTTCGCCCTGGATGCCAAA 3357  
Db 2873 GTGCTCGACGTTGTCACTGAAGCGGGAAGGACTGGCTGCTATTGGGCGAATATGCCGG 2932  
Qy 3358 CTCTGCTACCTGCTGGATGGAATCTCTTATCTATGTTGTCATTTCTCACTGCTCTTGTTC 3417  
Db 2933 CAGGATCTCTGTCATCTCACCTTGTCTCTCCGAGAAAGTATCCATCATGGCTGATGCA 2992  
Qy 3418 CTGAGAGTGAAGTTTCAAGCAGGCGCAGAGCCCCCGGTACCCAGAGGGCCAGAACAG 3477  
Db 2993 ATCGGGGCTGTCATACGCTTGATCCGCTACTGCTCCCATTTGACCAACAAGCGAACA 3052  
Qy 3478 CTCTATACGAGCTCAATCTAGGACGAGAGAGAGTACGATGTTTTGGACAAGAGAGT 3537  
Db 3053 CGC-----ATCGAGCGAGCACGTACTCGGATGGAAGCGGCTTTGTGATCAGGAT 3103  
Qy 3538 GGCCGGACCTGAGATGGGGGAAAGCCGAGAGGAAGAACCTTCAGAGAGGCTGTAC 3597  
Db 3104 GATCTGAGCAGAGCATCAGGG--GCTCGCGCCAGCGCAACTGTTGCCAGGCTCAAGG 3161  
Qy 3598 AATGAATGCAAGAAGATAAGATCGCGAGGCTTACAGTGAAGATTTGGGATGAAGGCGAG 3657  
Db 3162 CGCGCATGCCCGAGCGGAGGATCTGCTGCTGAGCCCATGGCGATGCTGCTTGGCGAATA 3221  
Qy 3658 CGCCGAGGGGCAAG--GGCACGATGGCTTTACAGGGTCTCAGTACAGCCACCAAGG 3715  
Db 3222 TCATGGTGGAAATGGCGCTTTTCTGGATTCATCGACTGTGGCGGCTGTGGTGGCGG 3281  
Qy 3716 ACACCTACGAGCGCTTCACATGAGCGCTTCCCTCGCTAA-----CTCGACGCGCC 3771  
Db 3282 ACCCTATCAGGACATAGCGTTGGCTACCGGTGATTTGCTGAAGAGCTTGGCGGGAAT 3341  
Qy 3772 GCGGATCCGGATTAGTCCAAATTTGTTAAAGACAGGATATCAGTGTCCAGGCTCTAGTTT 3831  
Db 3342 GGGCTGACCGCTTCTCTGCTTTACGCTATCGCGCTCCCGATTTCGACGCGATCGCT 3401  
Qy 3832 TGACTCAACATATCACAGCTGAAGCTATAGAGTAGGAGCCATAGATATAAATAAGA 3891  
Db 3402 TCTATCGCTTCTTTGACGAGTTCTTCTGAGCGGACTCTGGGGTTCGATAAATAAAGA 3461  
Qy 3892 TTTTATTTAGTCTCCAGAAAAGGGGAATGAAGACCCCACTGTAGTTTGGCAAGC 3951  
Db 3462 TTTTATTTAGTCTCCAGAAAAGGGGAATGAAGACCCCACTGTAGTTTGGCAAGC 3521  
Qy 3952 TAGCTTAAGTAACGCCATTTTGAAGCATGG--AAAATACATAAATCTGAGAATAGAGAAGT 4010  
Db 3522 TAGCTTAAGTAACGCCATTTTGAAGCATGGAATAATACATACTGAGATAGAGAAGT 3581  
Qy 4011 TCAGATCAAGGTTAGGAACAGA--GAGACAGAGAAATATGGCCCAACAGGATATCTTGG 4069  
Db 3582 TCAGATCAAGTCAAGAACAGATGGAACAGCTGAATATGGGCCAACAAGGATATCTTGG 3641  
Qy 4070 TAACAGTTCTCGTCCCG--GCTCAGGGCCCAAGACAGTTGGAACAGGAGAAATATGGCCAA 4128  
Db 3642 TAACAGTTCTCGTCCCGCTCAGGGCCCAAGACAGATGGAACAGTGAATATGGGCCAA 3701  
Qy 4129 ACAGGATATCTGTGTAAGCAGTTCTCGCCCGCTCAGGGCCCAAGAACAGATGTTCC 4188  
Db 3702 ACAGGATATCTGTGTAAGCAGTTCTCGCCCGCTCAGGGCCCAAGAACAGATGTTCC 3761  
Qy 4189 AGATGCGGTCCCGCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGGCCCA 4248  
Db 3762 AGATGCGGTCCAGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGGCCCA 3821



QY 4249 AGGACCTGAAATGACCCCTGTGCTTATTTTGAATACCAATCAGTTTCGCTTCTCGCTTCT 4308  
 DB 3822 AGACCTGAAATGACCCCTGTGCTTATTTTGAATACCAATCAGTTTCGCTTCTCGCTTCT 3881  
 QY 4309 GTTCCGCGCTTCTGTCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCCGCGCG 4368  
 DB 3882 GTTCCGCGCTTCTGTCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCCGCGCG 3941  
 QY 4369 CCAGTCTCCGATAGACTCGCTGCGCGGCTACCCGTGTTCTCAATATAACCCCTCTTGCGAG 4428  
 DB 3942 CCAGTCTCCGATAGACTCGCTGCGCGGCTACCCGTGTTCTCAATATAACCCCTCTTGCGAG 4001  
 QY 4429 TTGCAATCCGACTCGTGTCTCGCTGTTCTCGGAGGGTCT-CTCTGAGTGATTCAGTAC 4487  
 DB 4002 TTGCAATCCGACTCGTGTCTCGCTGTTCTCGGAGGGTCTCTCTGAGTGATTCAGTAC 4061  
 QY 4488 CCGTCAGCGGGGTCTTTTTCAGTTTCTCCACCTACACAGGTCTCACTAAATTCCTGATGT 4547  
 DB 4062 CCGTCAGCGGGGTCTTTTTCAGTTTGGGGTCTCGTCCGGATCGGAGACCCCTGCCAGGG 4121  
 QY 4548 GCGGAGGAGCTCCGTGACCCCGGTTTTTGTGTTATAATAAATGCAAGAACAGTGTCCC 4607  
 DB 4122 ACCACCGACCCACCGGAGGTAAAGCTG----- 4151  
 QY 4608 TTCAAGCCAGACTACATCTGACTCTCGCTTTATAAAGAAATGTTGAAGGCTCTGTGG 4667  
 DB 4152 -----GCTGCTCGCGGTTTCGGTGATGACGGTGAACCTCTG-AC 4193  
 QY 4668 ACTATCTGCACACGACTTTTTTAAGATTTTATATGCTCTCGATGAGGATTTAGTCAAT 4727  
 DB 4194 ACATGAGCTCCCGGAGCGGTACAGCTTGTCTGTAACGGATGCCGGGACAGACAAG 4253  
 QY 4728 CTATCTCTGTATTTTGTGCTGGTCTCGGTATTTTAAATTTTCTAGTTTGCACTCCCTTC 4787  
 DB 4254 CCGCTCAGGCGCGTCAAGCGGTTGTGGG----- 4283  
 QY 4788 CTGAGAGCAGCGGATGTCAGAGTAGTTAATACTCTGAGGGCAGGCTTCTGTGAAGGT 4847  
 DB 4284 ----- 4283  
 QY 4848 TGCCTGGGCTCAGTGTGAGATTTTGCCATAAAGGGTCTCTGCCCTGTGTACAGACAG 4907  
 DB 4284 ----- 4283  
 QY 4908 ATCGGAATCTAGAGTGCACTACAGTCCCGCGGTTCCGGGCTCTGATCTCAGGGCA 4967  
 DB 4284 -----GGTGTGCGGGCGCAGCCATGA----- 4304  
 QY 4968 TCTTTGCTTAGAGATCCTCTACCGCGGACGATCGTGGCGGGTACCGAGTTCGAATTGG 5027  
 DB 4305 ----- 4304  
 QY 5028 TAATCATGTGATAGTGTCTCTGTGTAATTTGTTATCCGCTCACAAATCCACACAAAC 5087  
 DB 4305 -----CCAGTCCAG 4314  
 QY 5088 ATACGAGCCGGAAGCAATAAGTGTAAAGCTGGGTGCTTAAGCTAGTGTAGTCACTACA 5147  
 DB 4315 TAGCGATAGCGGAGTGATCTAGTGGCTTAATCTAGGGCATCAGAGCAGATTTGATCTGAGA 4374  
 QY 5148 TTAATTTGCGTTCGGCTCACTGCGCGCTTTCCAGTCCGGGAAACCTGTCTGTCAGCTGCAT 5207  
 DB 4375 GT-----GCACATATGCGG 4389  
 QY 5208 TAATGAATCCGCAACGCGCGGAGAGCGGTTTTGCGTATTTGGGCGCTCTTCCGCTTCC 5267  
 DB 4390 TGTGAATACCGCAGACAGATGCGTAAGGAGAAATACCGCATCAGGGCTCTTCCGCTTCC 4449  
 QY 5268 TCGCTCACTGACTCGTCTGCGTCTCGGTGTTTCGGCTGCGGAGCGGTATCAGTCACTCA 5327  
 DB 4450 TCGCTCACTGACTCGTCTGCGTCTCGGTGTTTCGGCTGCGGAGCGGTATCAGTCACTCA 4509

QY 5328 AAGCGGTAATACGGTTATCCACAGATCAGGGATAACCGAGGAAAGCAATGTGAGCA 5387  
 DB 4510 AAGCGGTAATACGGTTATCCACAGATCAGGGATAACCGAGGAAAGCAATGTGAGCA 4569  
 QY 5388 AAGGCCAGCAAAAGGCCAGGAAACCGTAAAGGCGCGTTGCTGGCGTTTTTCCATAGG 5447  
 DB 4570 AAGGCCAGCAAAAGGCCAGGAAACCGTAAAGGCGCGTTGCTGGCGTTTTTCCATAGG 4629  
 QY 5448 CTCGCGCCCTTGACGAGCATCAAAAATCGACCTCAAGTCAAGGTGGGGAACCCG 5507  
 DB 4630 CTCGCGCCCTTGACGAGCATCAAAAATCGACCTCAAGTCAAGGTGGGGAACCCG 4689  
 QY 5508 ACAGGACTATAAAGATAACAGCGGTTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTGTT 5567  
 DB 4690 ACAGGACTATAAAGATAACAGCGGTTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTGTT 4749  
 QY 5568 CCGACCTTCCCGCTTACCGGATACCTGTCGCGCTTCTCCCTCGGGAAGCGTGGCTTT 5627  
 DB 4750 CCGACCTTCCCGCTTACCGGATACCTGTCGCGCTTCTCCCTCGGGAAGCGTGGCTTT 4809  
 QY 5628 TCTCATAGTCACTGCTGATGATCTCAGTTGCTGATGCTGCTTCCGCTCCAGCTGGG 5687  
 DB 4810 TCTCATAGTCACTGCTGATGATCTCAGTTGCTGATGCTGCTTCCGCTCCAGCTGGG 4869  
 QY 5688 TGTGTGACGAAACCCCGCTTCCAGCCCGCTGCGCTTATCCGCTAACTATCGTCTT 5747  
 DB 4870 TGTGTGACGAAACCCCGCTTCCAGCCCGCTGCGCTTATCCGCTAACTATCGTCTT 4929  
 QY 5748 GAGTCCAAACCCGGTAAGACAGACTTATCGGCACTGGCAGCAGCCACTGGTAAACAGGAT 5807  
 DB 4930 GAGTCCAAACCCGGTAAGACAGACTTATCGGCACTGGCAGCAGCCACTGGTAAACAGGAT 4989  
 QY 5808 AGCAGAGCGAGTATGTTAGGCGGTCTACAGAGTTCTTGAAGTGGTGGCTAACTACCG 5867  
 DB 4990 AGCAGAGCGAGTATGTTAGGCGGTCTACAGAGTTCTTGAAGTGGTGGCTAACTACCG 5049  
 QY 5868 TACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAATA 5927  
 DB 5050 TACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAATA 5109  
 QY 5928 AGAGTTGGTAGTCTTTGATCCGGCAAAACAAACCCGCTGCTGAGCGGTGGTTTTTGT 5987  
 DB 5110 AGAGTTGGTAGTCTTTGATCCGGCAAAACAAACCCGCTGCTGAGCGGTGGTTTTTGT 5169  
 QY 5988 TCCAGACGACAGATTTACCGCGAGAAAGATCTCAAGAGATCTTGTATCTTTCT 6047  
 DB 5170 TCCAGACGACAGATTTACCGCGAGAAAGATCTCAAGAGATCTTGTATCTTTCT 5229  
 QY 6048 ACGGGTCTGACGCTCAGTGGAAACGAAACCTCAGCTTAAGGATTTTGGTCAATGAGTTA 6107  
 DB 5230 ACGGGTCTGACGCTCAGTGGAAACGAAACCTCAGCTTAAGGATTTTGGTCAATGAGTTA 5289  
 QY 6108 TCAAAAAGGATCTTCACTTAGATCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAA 6167  
 DB 5290 TCAAAAAGGATCTTCACTTAGATCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAA 5349  
 QY 6168 AGTATATAGTAACTTGTCTGACAGTTACCAATGCTTAATCAGTGAGGACCTATC 6227  
 DB 5350 AGTATATAGTAACTTGTCTGACAGTTACCAATGCTTAATCAGTGAGGACCTATC 5409  
 QY 6228 TCAGGATCTGTCTATTTTCTGCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6287  
 DB 5410 TCAGGATCTGTCTATTTTCTGCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5469  
 QY 6288 ACATACGAGGAGGCTTACCATCTGCGCCCACTGCTGCAATGATACCGAGACCCACG 6347  
 DB 5470 ACATACGAGGAGGCTTACCATCTGCGCCCACTGCTGCAATGATACCGAGACCCACG 5529  
 QY 6348 TCACCGGCTCCAGATTTATCAGCAATAAACCGAGCCGAGGAGGCGGAGCGGAGGAGT 6407  
 DB 5530 TCACCGGCTCCAGATTTATCAGCAATAAACCGAGCCGAGGAGGCGGAGGAGGAGT 5589  
 QY 6408 GGTCTCTCAACTTTTATCGGCTCCATCCAGTCTATTAATTTGTTGCGGGGAAGCTAGAGTA 6467

```

Db 5590 GGTCTGCACTTTATCCGCTCCATCAGTCTTAAATTTGTCGGGAAGCTAGAGTA 5649
Qy 6468 AGTAGTTCGCCAGTAAATAGTTTGGCAACGTTGTGTCATTTGCTACAGGC-TGTTGGTG 6526
Db 5650 AGTAGTTCGCCAGTAAATAGTTTGGCAACGTTGTGTCATTTGCTGCAAGCATCGTGGTG 5709
Qy 6527 TCAGCTCGTGGTGGTATGGCTTCATCAGTCCGTTCCCAACGATCAAGGGAGTT 6586
Db 5710 TCAGCTCGTGGTGGTATGGCTTCATCAGTCCGTTCCCAACGATCAAGGGAGTT 5769
Qy 6587 ACATGATCCCATGCTGTGCAAAAAGCGTTAGTCTCTTCGGTCCCTCCGATCGTTGTC 6646
Db 5770 ACATGATCCCATGCTGTGCAAAAAGCGTTAGTCTCTTCGGTCCCTCCGATCGTTGTC 5829
Qy 6647 AGAAGTAAAGTTGGCCGAGTGTATCACTCATGTTATGGCAGCAGTGCATAATTTCTTT 6706
Db 5830 AGAAGTAAAGTTGGCCGAGTGTATCACTCATGTTATGGCAGCAGTGCATAATTTCTTT 5889
Qy 6707 ACTGTATGCCATCCGTAAGATGCTTTTCTGTGACTGTGTAGTACTCAACCAAGTCAATC 6766
Db 5890 ACTGTATGCCATCCGTAAGATGCTTTTCTGTGACTGTGTAGTACTCAACCAAGTCAATC 5949
Qy 6767 TGAGAAATAGTGTATGCGCGGAGCGAGTTGCTTTCGCCGCGTCAATACGGGATAATACC 6826
Db 5950 TGAGAAATAGTGTATGCGCGGAGCGAGTTGCTTTCGCCGCGTCAATACGGGATAATACC 6009
Qy 6827 GCGCCACATAGCAGAACTTTAAAGTGTCTCATTTGGAACACGTTCTTCGGGGCGAAAA 6886
Db 6010 GCGCCACATAGCAGAACTTTAAAGTGTCTCATTTGGAACACGTTCTTCGGGGCGAAAA 6069
Qy 6887 CTCTCAAGGATCTTACCGCTGTGAGATGCCAGTTGATGTAACCCACTCGTGCACCCAAAC 6946
Db 6070 CTCTCAAGGATCTTACCGCTGTGAGATGCCAGTTGATGTAACCCACTCGTGCACCCAAAC 6129
Qy 6947 TGATCTTCAGCATCTTTTACTTTTACACAGCGTTTCTGGGTGAGCAAAACAGGAGGCAA 7006
Db 6130 TGATCTTCAGCATCTTTTACTTTTACACAGCGTTTCTGGGTGAGCAAAACAGGAGGCAA 6189
Qy 7007 AATCGCGCAAAAAGGGAATAAGGCGCACACGGAATGTTGAATACTCATACTCTTCTCTT 7066
Db 6190 AATCGCGCAAAAAGGGAATAAGGCGCACACGGAATGTTGAATACTCATACTCTTCTCTT 6249
Qy 7067 TTCAATATTTAAGCAATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAA 7126
Db 6250 TTCAATATTTAAGCAATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAA 6309
Qy 7127 TGTATTTAGAAAATAAACAATAGGAGTTTCGGCGCACATTTCCCGGAAAAGTGCACCT 7186
Db 6310 TGTATTTAGAAAATAAACAATAGGAGTTTCGGCGCACATTTCCCGGAAAAGTGCACCT 6369
Qy 7187 GACGCTTAAGAAACCATTTATCATGACATTTAACTTATAAATAAGGCGTATCAGGAGG 7246
Db 6370 GACGCTTAAGAAACCATTTATCATGACATTTAACTTATAAATAAGGCGTATCAGGAGG 6429
Qy 7247 CCCTTTCGTCT 7257
Db 6430 CCCTTTCGTCT 6440

```

RESULT 11  
 US-09-963-206B-4  
 ; Sequence 4, Application US/09963206B  
 ; Patent No. US20020123076A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ferrick, David A.  
 ; APPLICANT: Swift, Susan E.  
 ; APPLICANT: Armstrong, Randall  
 ; APPLICANT: Fox, Bryan  
 ; TITLE OF INVENTION: Methods and Compositions for Screening for Modulators and Ige Syn  
 ; TITLE OF INVENTION: Secretion and Switch Rearrangement  
 ; FILE REFERENCE: A-66038-3/RMS/JUD/DLR  
 ; CURRENT APPLICATION NUMBER: US/09/963.206B

```

; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 09/076,624
; PRIOR FILING DATE: 1998-05-12
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 6219
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-963-206B-4

```

Query Match 40.3%; Score 3082.2; DB 9; Length 6219;  
 Best Local Similarity 73.7%; Pred. No. 0;  
 Matches 4587; Conservative 0; Mismatches 958; Indels 680; Gaps 25;

```

Qy 958 CCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGGCGCCAGTCTCCGATAGACTG 1017
Db 662 CAGAGCTCAATAAAGAGCCCAACCCCTCACTCGGGCGCCAGTCTCCGATAGACTG 721
Qy 1018 CGTCCCGGGGTACCCGTTATCCCAATAAAGCCCTCTTGTCTTTGCAATCCGAATCGTGA 1077
Db 722 AGTCCCGGGGTACCCGTTATCCCAATAAAGCCCTCTTGCAG-TTGCATCCGACTTGTGT 780
Qy 1078 CTGCTGATCTTGGAGGGTCTCTCAGATTGATGACTGCC-ACCTCGGGGTCTTT 1136
Db 781 CTGCTGTTCTTGGAGGGTCTCTCTGAGTGTGACTACCCGTACAGCGGGGTCTTT 840
Qy 1137 CATTTGGAGTTCCACCGAGATTTGGAGACCCCTCCCGAGGACCCGACCCCGCC 1196
Db 841 CATTTGGGGGTCTGTCGGGATCGGGAGACCCCTGCCAGGGACCCCGA-CCCAACCACC 899
Qy 1197 GCGAGTTAAGCTGCCAGCAACTTATCTGTCTGTCCGATTGTCTAGTGTCTATGACTG 1256
Db 900 GCGAGTTAAGCTGCCAGCAACTTATCTGTCTGTCCGATTGTCTAGTGTCTATGACTG 959
Qy 1257 ATTTTATGCGCTCGTCCGTTAGTTAGTAACTAGTCTGTATCTGCGGACCCGTCG 1316
Db 960 ATTTTATGCGCTCGTCCGTTAGTTAGTAACTAGTCTGTATCTGCGGACCCGTCG 1019
Qy 1317 GTGAACTGACGAGTTTCGGAACACCCCGCCGCAACCTCGGAGACGTCGAGGACCTTCG 1376
Db 1020 GTGAACTGACGAGTTTCGGAACACCCCGCCGCAACCTCGGAGACGTCGAGGACCTTCG 1079
Qy 1377 GGGCCCGTTTTTGGCCCGGACCTGAGTCTTAAATCCGATCCGATCTTTAGGACTCTTGGT 1436
Db 1080 GGGCCCGTTTTTGGCCCGGACCTGAGTCCAAAATCCGATCCGATCTTTGGACTCTTGGT 1139
Qy 1437 GCACCCCGCTTAGAGGAGGATATGTGGTTCTGGTAGGAGACGAGAACCTAAAAACAGTTC 1496
Db 1140 GCACCCCGCTTAGAGGAGGATATGTGGTTCTGGTAGGAGACGAGAACCTAAAAACAGTTC 1199
Qy 1497 CCGCTCCGTCTGAATTTTGTCTTTGGGACCGAAGCCGCGCGCGCTCTTGTCT 1556
Db 1200 CCGCTCCGTCTGAATTTTGTCTTTGGGACCGAAGCCGCGCGCGCTCTTGTCT 1259
Qy 1557 TGCTGACGATCGTTCTGTGTCTGTCTGCTGACTGTGTTTCTGTTATTTCTGTAATAAT 1616
Db 1260 TGCTGACGATCGTTCTGTGTGTCTGTCTGCTGACTGTGTTTCTGTTATTTCTGTAATAAT 1319
Qy 1617 ATGGGCGCGGCTAGACTGTTTACCACCTCCCTTAAAGTTTGACTTTAGCTTACTGGAAAGAT 1676
Db 1320 ATGGGCGCGGCTAGACTGTTTACCACCTCCCTTAAAGTTTGACTTTAGCTTACTGGAAAGAT 1379
Qy 1677 GTCAGCGGATCGCTCAACAGTCTGAGTGTCTCAAGAGAGAGCGTTGGGTACTCTTC 1736
Db 1380 GTCAGCGGATCGCTCAACAGTCTGAGTGTCTCAAGAGAGAGCGTTGGGTACTCTTC 1439
Qy 1737 TGCTCTCAGAAATGGCCAACTTTAAGCTGCGATGGCGCGGAGACGACCTTTAAACCGA 1796
Db 1440 TGCTCTCAGAAATGGCCAACTTTAAGCTGCGATGGCGCGGAGACGACCTTTAAACCGA 1499

```



Db 3554 AGCTTAAGTTAAGCCATTTTTCAGAGCATGGAATAATCATTAACATGAGATAGAGAAGTT 3613  
Qy 4012 CAGATCAAGGTTAGAAACAGA-GAGACAGCAGAAATATGGGCCAAACAGGATATCTGTGT 4070  
Db 3614 CAGATCAAGGTCAGGAAACAGATGGAACAGCTGAAATATGGGCCAAACAGGATATCTGTGT 3673  
Qy 4071 AAGCAGTTCTTCCGCC- GCTCAGGGCCAGAACAGTTCGGAACAGAGAAATATGGGCCAAA 4129  
Db 3674 AAGCAGTTCTTCCGCCCTCAGGGCCAGAACAGATGGAACAGCTGGAATATGGGCCAAA 3733  
Qy 4130 CAGGATATCTGTGTGAAGCAGTTCTGCCCGGCTCAGGGCCAGAAACAGATGTCGCCA 4189  
Db 3734 CAGGATATCTGTGTGAAGCAGTTCTGCCCGGCTCAGGGCCAGAAACAGATGTCGCCA 3793  
Qy 4190 GATCGGTCCTCCGCTCAGCAGTTCTTAGAGAACCATCAGATGTTTCCAGGGTGCCTCAA 4249  
Db 3794 GATCGGTCCTCCGCTCAGCAGTTCTTAGAGAACCATCAGATGTTTCCAGGGTGCCTCAA 3853  
Qy 4250 GGACCTGAATGACCTGTGCTTATTTGAACCTAACCAATCAGTTCGCTTCTCGCTTCTG 4309  
Db 3854 GGACCTGAATGACCTGTGCTTATTTGAACCTAACCAATCAGTTCGCTTCTCGCTTCTG 3913  
Qy 4310 TTCCGGCGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCCGGCGGC 4369  
Db 3914 TTCCGGCGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCCGGCGGC 3973  
Qy 4370 CAGTCTCCGATAGACTCGTCCCGGGTACCCGTGTTCTCAATAAACCCCTCTTGCACT 4429  
Db 3974 CAGTCTCCGATAGACTCGTCCCGGGTACCCGTGTTCTCAATAAACCCCTCTTGCACT 4033  
Qy 4430 TGCATCCGACTCGTGTCTCGCTGTTCTTGGAGGGTCT-CTCTGAGTGAATGACTACC 4488  
Db 4034 TGCATCCGACTCGTGTCTCGCTGTTCTTGGAGGGTCTCTCTGAGTGAATGACTACC 4093  
Qy 4489 CGTCAGCGGGTCTTTCAGTTTCTCCCACTACACAGGCTCTCAACAATTCCTGATGTG 4548  
Db 4094 CGTCAGCGGGTCTTTCAGTTTCTCCCACTACACAGGCTCTCAACAATTCCTGATGTG 4113  
Qy 4549 CCGCAGGAGCTCCGTCAGCCCGGTTTGTGTTTATAAATAAAGCAAGTGTCCCT 4608  
Db 4114 ----- 4113  
Qy 4609 TCAAGCCAGACTACATCTGACTCTCGGCTTTATAAAGAATGTGAAGGCTCTGTGGA 4668  
Db 4114 ----- 4113  
Qy 4669 CTATCTGCCACAGACTTTTAAAGATTTTATGCTCTCGATGAGGATTTAGTCAATC 4728  
Db 4114 ----- 4113  
Qy 4729 TATCCTCGTCTATTTTGTGCTTCTCGGTATTTTAAATTTCTAGTTTGCACCTCCCTTCC 4788  
Db 4114 ----- 4113  
Qy 4789 TGAGAGCAGCGGATTCAGAGTAGTTAATACTCTGAGGGCAGGCTTCTGTGAAGGTT 4848  
Db 4114 -----TCCGACTGTGTGCTCGCTCTGGGA 4141  
Qy 4849 GCCTGGGCTCAGTGTGAGATTTGGCAATAAAGGGGTCCTGCCCTGTGTACAGACAGA 4908  
Db 4142 GGGTCTCCTCTGAGTGTGACTACCCGTCAGCGGGGCTCTTCACATG----- 4189  
Qy 4909 TCGGAATCTAGAGTGCACTACAGAGTCCCGCGGTTCCGGGGCTCTGATCTCAGGGCAT 4968  
Db 4190 ----- 4189  
Qy 4969 CTTTGCTTAGAGATCCTACCGCGAGCATCGTGGCCGGGTACCGAGCTCGAATTCGT 5028  
Db 4190 ----- 4189  
Qy 5029 AATCATGTCTAGTGTTCCTGTGTGAATTTGTTATCCGCTCACAATTCACACACA 5088  
Db 4190 -----CAGCATGTATCAAAAATTAAATTTGGTTTTTTTCTTAAAGTATTACATTA 4238

Qy 5089 TACGAGCGGGAAGCATAAAGTGTAAAGCCTCGGGTGCCTAATGAGTGAAGTCAACTCACAT 5148  
Db 4239 AA----- 4240  
Qy 5149 TAATTTGGCTTGGCTCACTGCCCCCTTTTCAGTCCGGGAAACCTGTGTCGCCAGCTGCATT 5208  
Db 4241 -----TGGCCATAGTTGCATT 4256  
Qy 5209 AATGAATCGGCCAACCGCGGGGAGAGCGGTTTTCGTAATTTGGGCGCTCTTCCGCTTCT 5268  
Db 4257 AATGAATCGGCCAACCGCGGGGAGAGCGGTTTTCGTAATTTGGGCGCTCTTCCGCTTCT 4315  
Qy 5269 CGCTCACTGACTCGCTCGCTTCGTTTCGCTTCGCGCAGCGGTATCAGCTCACTCAA 5328  
Db 4316 CGCTCACTGACTCGCTCGCTTCGTTTCGCTTCGCGCAGCGGTATCAGCTCACTCAA 4375  
Qy 5329 AGGCGGTAATACGGTTATCCACAGANTCAGGGGATACCGCAGGAAGAAACATGTGAGCAA 5388  
Db 4376 AGGCGGTAATACGGTTATCCACAGANTCAGGGGATACCGCAGGAAGAAACATGTGAGCAA 4435  
Qy 5389 AAGCCAGCAAAAAGCCAGAGAACCGTAAAGAGCGGTTCTGCTGGCGTTTTTCCATAGGC 5448  
Db 4436 AAGCCAGCAAAAAGCCAGAGAACCGTAAAGAGCGGTTCTGCTGGCGTTTTTCCATAGGC 4495  
Qy 5449 TCCGCCCTCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGA 5508  
Db 4496 TCCGCCCTCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGA 4555  
Qy 5509 CAGGACTATTAAGATACAGGCGTTTCCCTCGGAAAGTCCCTCGTGGGCTCTCTCTGTTTC 5568  
Db 4556 CAGGACTATTAAGATACAGGCGTTTCCCTCGGAAAGTCCCTCGTGGGCTCTCTCTGTTTC 4615  
Qy 5569 CGACCTGCGGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAGCGTGGCGTTT 5628  
Db 4616 CGACCTGCGGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAGCGTGGCGTTT 4675  
Qy 5629 CTCACTAGCTCAGCTGTAGGTATCTCAGTTCCGTTGTAGGTCTGCTCCAGCTCCAGCTGGGCT 5688  
Db 4676 CTCACTAGCTCAGCTGTAGGTATCTCAGTTCCGTTGTAGGTCTGCTCCAGCTGGGCT 4735  
Qy 5689 GTGTGCAAGAACCCCGCTTCAGCCGACGCTCGGCTTATCCGGTAACTATCTGCTCTTG 5748  
Db 4736 GTGTGCAAGAACCCCGCTTCAGCCGACGCTCGGCTTATCCGGTAACTATCTGCTCTTG 4795  
Qy 5749 AGTCCAAACCCGGTAAGACACGACTTATCGCACCTGGCAGCAGCCACTGGTAAACAGGATTA 5808  
Db 4796 AGTCCAAACCCGGTAAGACACGACTTATCGCACCTGGCAGCAGCCACTGGTAAACAGGATTA 4855  
Qy 5809 GCAGAGCGAGGTATGTAGGGGCTGTACAGAGTTCTTGAAGTGGTGGCTTAACCTACGGCT 5868  
Db 4856 GCAGAGCGAGGTATGTAGGGGCTGTACAGAGTTCTTGAAGTGGTGGCTTAACCTACGGCT 4915  
Qy 5869 ACCTAGAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAA 5928  
Db 4916 ACCTAGAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAA 4975  
Qy 5929 GAGTTGTAGCTCTTGTATCCGGCAACAAACACCGCTGTGTAGCGGTGGTTTTTGTGTTT 5988  
Db 4976 GAGTTGTAGCTCTTGTATCCGGCAACAAACACCGCTGTGTAGCGGTGGTTTTTGTGTTT 5035  
Qy 5989 GCAAGCAGCAGATTACCGCGCAAAAAAAGGATCTCAAGAGATCTCTTGTATCTTTTCTTA 6048  
Db 5036 GCAAGCAGCAGATTACCGCGCAAAAAAAGGATCTCAAGAGATCTCTTGTATCTTTTCTTA 5095  
Qy 6049 CGGGGTCTGACGCTCAGTGGAAACGAAACTCACTTAAGGGATTTGGTCAATGAGATTAT 6108  
Db 5096 CGGGGTCTGACGCTCAGTGGAAACGAAACTCACTTAAGGGATTTGGTCAATGAGATTAT 5155  
Qy 6109 CAAAAAGGATCTTCCACCTAGATCTTTTAAATTTAAATTTAAAGTTT---TAAATCAATCT 6165  
Db 5156 CAAAAAGGATCTTCCACCTAGATCTTTTAAATTTAAATTTAAAGTTTGGCAATCAATCT 5215





Qy 1677 GTCGAGCGGATGCTCAAAACAGTCGGTAGATGTCAAAGAGAGACGTTGGTTACCTTC 1736  
 Db |||||  
 Qy 1737 TGCTCTGAGAAATGCGCAACCTTTAACTCGGATGCGCGGAGCGGCACTTTAAACCGA 1796  
 Db |||||  
 Qy 1797 GACCTCATCACCAGGTTAAGATCAAGGTCTTTTCACTGCGCGCGCATGGACACCCAGAC 1856  
 Db |||||  
 Qy 1857 CAGGTCCCTTACATCGTGAACCTTGGGAAGCTTTGGCTTTTGAACCCCTTCCCTGGGTCAAG 1916  
 Db |||||  
 Qy 1917 CCCTTGTGACACCTTAAGCCTCCGCTCCTCTCTCTCCATCGCGCGGTCTCTCCCTTT 1976  
 Db |||||  
 Qy 1977 GAACCTCTCGTTGACCCCGCTCGATCTCCCTTTATCCAGCCCTCACTCCTTCTCTA 2036  
 Db |||||  
 Qy 2037 GCGCGCCCATATGCGCCATATGAGATCTTATATGGGCAACCCCGCCCTTTGTAACCTTC 2096  
 Db |||||  
 Qy 2097 CCGTACCTGATGACAGAGATTACTAAAGCCCTCTCTCCAGCCCTCACTTACAGGCT 2156  
 Db |||||  
 Qy 2157 -TCTACTTGTAGTCCAGACGAACTGTGAGACTCTGGCGGAGCTTCAAGAAACAACCTG 2215  
 Db |||||  
 Qy 2216 GACCGACGGTGTACTTACCTTACCGAGTTCGGCGGACACAGTGTGGTTCGCGGACAC 2275  
 Db |||||  
 Qy 2276 CAGACTAAGAACCTTAGAACCTCGCTGGAAGAGACTTACAGAGTCTGTGACACCCCTC 2335  
 Db |||||  
 Qy 2336 ACCGCCCTCAAAGTAGAGCGCATCGAGCTTGGATACACGCCCGCCAGTGAAGCTGCC 2395  
 Db |||||  
 Qy 2395 GACCGCGGGTGGACCATCTCTAGACTGCC-----GGATCTCGAGGG 2455  
 Db |||||  
 Qy 2456 TGGTAGCAACAGCTACAGGTGTCACTCCGATCTCCAGCTGACCCAGAGCCAGCAGCC 2515  
 Db |||||  
 Qy 2515 TGAGCGCAGCGTGGGTGACAGGTGACCATCACTGTAAAGCCAGTCAGATGTGGTA 2575  
 Db |||||  
 Qy 2575 TGAGCGCAGCGTGGGTGACAGGTGACCATCACTGTAAAGCCAGTCAGATGTGGTA 2635  
 Db |||||  
 Qy 2635 CATCACCCGCAACACTGGTGTGCAAGAGATTCAAGCGTGAAGCGGTAGCGGTACCGACT 2695  
 Db |||||  
 Qy 2695 TCACCTTACCATCAGCAGCTTCCAGCCAGAGGACATCGCCACCTACTACTTGCAGCAAT 2755  
 Db |||||  
 Qy 2755 GCAAGCGCGATCA-----CATGGTCTCTGGAGTTCGTGACCGCGCGCGGAT 3433  
 Db |||||  
 Qy 3433 GACTCAACAATATACCCAGCTGAAGCCCTATAGAGTAGAGCAGCCATAGATAAATAAAGAT 3892  
 Db |||||

Qy 2756 ATAGCCTCTATCGGTCTGTTGCGCCCAAGGGACCAAGGTGGAATCAAACGAGGTGCTCAG 2815  
 Db |||||  
 Qy 2815 GATCGGTGGATCGGCTCTGCTCAGGATCGGAGGTCCAACTGGTGGAGAGCGGTG 2875  
 Db |||||  
 Qy 2875 GAGGTGTTGTGCAACCTGCGCGCTCCCTGCGCCCTGTCTGCTCGGCATCTGGCTTCGATT 2935  
 Db |||||  
 Qy 2935 TGAGTTGGATAGTTGTGGAAGAGTCAAAATGGCTCTCTCTCAAGCGTATTCAACAGGGGC 2971  
 Db |||||  
 Qy 2971 TCACCAATATTTGATGAGTTGGGTGAGACAGGCACTGGAAGAGTTCTTGAAGTGAATG 2995  
 Db |||||  
 Qy 2995 TGAAGGATGCGCAAGAGTACCCCATTTGATGAGGATCTGATCTGGGCTCGGTGCACAT 2631  
 Db |||||  
 Qy 2631 GAGAAATTCATCCAGATAGCAGTACGATTAACTATGCGCGCTCTCTAAAGATAGATTTA 3055  
 Db |||||  
 Qy 3055 GCTTTACATGTTGTTAGTTCGAGGTTAAACACGCTAGGCCCCCGAACCACGGGACGT 2691  
 Db |||||  
 Qy 2691 CAATATCGGAGACACGCCCAAGAAACATTTGTTCTCTGMAATGGACAGCCTGAGACCG 3115  
 Db |||||  
 Qy 3115 GGTTTCTCTTTGAAAAACAGATGATATATG-----GGGATCCACCGG 2736  
 Db |||||  
 Qy 2736 AAGACACGGGTCTATTTTGTGCAAGCCTTTACTTGGCTTCCCTGGTTGCTTAT 3175  
 Db |||||  
 Qy 3175 TCGCACCATGTTGAGCAAGGCGGAGGAGCTGTTTACCGGGTG--GTGCCATCTCGT 2794  
 Db |||||  
 Qy 2794 GGGGCCAAGGACCCCGGTCACTGCTCCAGTGTAAAGCCACCAAGCGGAGGCGGCGC 3235  
 Db |||||  
 Qy 3235 CGAGCTGGAGCGGACGTAACCGCCCAAGTTTCCAGCGTGTCCGCGGAGGCGGAGGCGA 2854  
 Db |||||  
 Qy 2854 GACCACCAACACCGCGCCCAACCATCGCTCGCAGCCCTCTGCTGCGCCAGAGGGCG 3295  
 Db |||||  
 Qy 3295 TGCCACCTACGGCAA--GCTGACCTGAGTTTCTGCAACACCGGAGTGTCCCGTGC 2913  
 Db |||||  
 Qy 2913 CTGCGCAGCGGCGGGGCGCAGTGCACAGAGGGGTGAGATTGCGCCCTGGATCCCA 3355  
 Db |||||  
 Qy 3355 CTTGGCCACCTCGTGACACACCTGACCTAGCGCGTGCAGTGTCTTACGCGCTACCCCG 2973  
 Db |||||  
 Qy 2973 AACTCTGCTACTGCTGGATGGAATCTCTTCTCATCTATGGTGTCTATCTACTGCTTGT 3415  
 Db |||||  
 Qy 3415 ACCACATGAGACGACGAA-----CTTCTTCAAGTCCGCCATGCCGAAGGCTACG 3024  
 Db |||||  
 Qy 3024 TCTTGTAGAGTGAAGTTTCAAGAGGAGCGCAGAGCCCGCGCTACAGCAGGCGCAGAAC 3475  
 Db |||||  
 Qy 3475 TCCAGGAGCGCAACCATTTCTTCAAGGACGAGCGCAACTACAGACCGCGCCGAGGTGA 3084  
 Db |||||  
 Qy 3084 AGCTCTATAACGAGCTCAATCTTAGGACGAAAGAGGAGTAGATGATGTTTGGACAGAGAC 3535  
 Db |||||  
 Qy 3535 AGTTGAGGGCGACACCTGTTGAACCGCATCGAGCTGAAGGCATCGACTTCAAGGAGG 3144  
 Db |||||  
 Qy 3144 GTGGCCGGAGCCCTGAGATGGGGGAAAGCGGAGAGAGAAACCTCAGG---AAGGCC 3592  
 Db |||||  
 Qy 3592 ACAGCAACATCTCTGGGCGCAAGCTGGAGTACACTACAAACAGCCACAACTCTATATCA 3204  
 Db |||||  
 Qy 3204 TGTACATGATGATGCAAAAGATAAGATGCGGAGGCGCTACAGTGAGATTGGGATGAAG 3652  
 Db |||||  
 Qy 3652 TGGCCGCAAGAGAGACGCGCATCAAGGTGAATCTTCAAGATCCGCCCAACATCTGAGG 3264  
 Db |||||  
 Qy 3264 GCGAGCCCGGAGGGCAAGGGCGACGATGGCTTTTACAGGGGTCTCAGTACAGCCACCA 3712  
 Db |||||  
 Qy 3712 ACAGCAGCTGAGCTCGCGACCACTTACAGCAGAACACCCCATCGGCGAGCGGCCCG 3324  
 Db |||||  
 Qy 3324 AGGACACTTACGAGCGCTTCAATGAGCGCTTGGCCCTCGCTAACTCGAGCGGGCGG 3772  
 Db |||||  
 Qy 3772 TGCTGTGCGCGCAACCACTACCTGAGCACCCAGCTCGGCTCGGCTGAGCAAGACCCCAACG 3384  
 Db |||||  
 Qy 3384 CGGATCCGGATTAGTCCAAATTTGTTAAAGACAGATATCAGTGTGCTCAGGCTCTAGTTTT 3832  
 Db |||||  
 Qy 3832 AGAAGCGCGATCA-----CATGGTCTCTGGAGTTCGTGACCGCGCGCGGAT 3433  
 Db |||||  
 Qy 3433 GACTCAACAATATACCCAGCTGAAGCCCTATAGAGTAGAGCAGCCATAGATAAATAAAGAT 3892  
 Db |||||





```
QY 6049 CCGGCTCTGACGCTCAGTGGGAAACGAAACTCAGCTTAAGGATTTTGGTCTATGAGATTAT 6108
Db 5096 CCGGCTCTGACGCTCAGTGGGAAACGAAACTCAGCTTAAGGATTTTGGTCTATGAGATTAT 5155
QY 6109 CAAAAGGATCTTACACCTAGATCCTTTTAAATTAATAATGAAGTTT---TAAATCAATCT 6165
Db 5156 CAAAAGGATCTTACACCTAGATCCTTTTAAATTAATAATGAAGTTTTCGGCAATCAATCT 5215
QY 6166 AAAGTATATAGTAAGTAACCTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGACACCTA 6225
Db 5216 AAAGTATATAGTAAGTAACCTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGACACCTA 5275
QY 6226 TCTCAGCATCTGTCTATTGCTTCATCCATAGTTTGCTGACTCCCGCTGCTGTAGATAA 6285
Db 5276 TCTCAGCATCTGTCTATTGCTTCATCCATAGTTTGCTGACTCCCGCTGCTGTAGATAA 5335
QY 6286 CTAGATACGGAGGGCTTACCATCTGCGCCAGTGTGCAATGATATGATACCGGAGCCAC 6345
Db 5336 CTAGATACGGAGGGCTTACCATCTGCGCCAGTGTGCAATGATATGATACCGGAGCCAC 5395
QY 6346 GCTCACCCTCCAGATTTATCAGCAATAAACCCAGCCAGCGGAGGCGGAGGAGAA 6405
Db 5396 GCTCACCCTCCAGATTTATCAGCAATAAACCCAGCCAGCGGAGGCGGAGGAGAA 5455
QY 6406 GTGGTCTGCAACTTTATCCGCTCCATCCAGTCTTAAATTTGTCGGGAAAGCTAGAG 6465
Db 5456 GTGGTCTGCAACTTTATCCGCTCCATCCAGTCTTAAATTTGTCGGGAAAGCTAGAG 5515
QY 6466 TAAGTAGTTCGCGAGTTAATAGTTTGGCGCAAGTTTGGCATTTGCTACAGGC-TCGTGG 6524
Db 5516 TAAGTAGTTCGCGAGTTAATAGTTTGGCGCAAGTTTGGCATTTGCTACAGGCATCGTGG 5575
QY 6525 TGTCAAGCTCGTGGTTGGTATGGCTTCATTTCAGCTCCGGTTCCCAAGATCAAGGCGAG 6584
Db 5576 TGTCAAGCTCGTGGTTGGTATGGCTTCATTTCAGCTCCGGTTCCCAAGATCAAGGCGAG 5635
QY 6585 TTACATGATCCCCATGTTGTGCAAAAAGCGGTTAGTCTCTCGTCCCTCCGATCGTTG 6644
Db 5636 TTACATGATCCCCATGTTGTGCAAAAAGCGGTTAGTCTCTCGTCCCTCCGATCGTTG 5695
QY 6645 TCAGAAGTAAGTTGGCCGAGTGTATACATCATGTTTATGGCAGCATGCAATATCTC 6704
Db 5696 TCAGAAGTAAGTTGGCCGAGTGTATACATCATGTTTATGGCAGCATGCAATATCTC 5755
QY 6705 TTACTGTTCATGCAATCCGTAAGATGCTTTCTGTGACTGGTGAGTACTCAACCAAGTCAT 6764
Db 5756 TTACTGTTCATGCAATCCGTAAGATGCTTTCTGTGACTGGTGAGTACTCAACCAAGTCAT 5815
QY 6765 TCTGAGATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATA 6824
Db 5816 TCTGAGATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATA 5875
QY 6825 CCGGCGCACATAGCAGAACTTTAAAGTGTATCATATTTGGAATAAGTTTCTCGGCGGAA 6884
Db 5876 CCGGCGCACATAGCAGAACTTTAAAGTGTATCATATTTGGAATAAGTTTCTCGGCGGAA 5935
QY 6885 AACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTGAGTAAACCACTCGTGCACCCA 6944
Db 5936 AACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTGAGTAAACCACTCGTGCACCCA 5995
QY 6945 ACTGATCTCAGCATCTTTTAACTTTTACAGCGGTTTCTGGGTGAGCAAAAACAGAGGC 7004
Db 5996 ACTGATCTCAGCATCTTTTAACTTTTACAGCGGTTTCTGGGTGAGCAAAAACAGAGGC 6055
QY 7005 AAAATGCCGCAAAAAGGGAATAAGGCGGACACGGAATTTGGAATATCTACTCTTCC 7064
Db 6056 AAAATGCCGCAAAAAGGGAATAAGGCGGACACGGAATTTGGAATATCTACTCTTCC 6115
QY 7065 TTTTTCATATATTTGAAGCAATTTATCAGGGTTATTTGTTCTCATGAGCGGATACATTTG 7124
Db 6116 TTTTTCATATATTTGAAGCAATTTATCA-GGTTATTTGTTCTCATGAGCGGATACATTTG 6174
```

```
QY 7125 AATGTATTAGAAAAATAAACAATAGGGTTTCGGCGCACATTTTC 7169
Db 6175 AATGTATTAGAAAAATAAACAATAGGGTTTCGGCGCACATTTTC 6219
```

## RESULT 13

```
US-09-963-247A-4
; Sequence 4, Application US/09963247A
; Publication No. US20040214155A1
; GENERAL INFORMATION:
; APPLICANT: Ferrick, David A.
; APPLICANT: Swift, Susan E.
; APPLICANT: Armstrong, Randall
; APPLICANT: Fox, Bryan
; TITLE OF INVENTION: Methods and Compositions for Screening for Modulators and Ige Syn
; FILE OF INVENTION: Secretion and Switch Rearrangement
; FILE REFERENCE: A-66038-1/RMS/JJD/DLR
; CURRENT APPLICATION NUMBER: US/09/963,247A
; CURRENT FILING DATE: 2002-05-07
; PRIOR FILING DATE: 1998-05-12
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 6219
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-963-247A-4
```

```
Query Match 40.3%; Score 3082.2; DB 11; Length 6219;
Best Local Similarity 73.7%; Pred. No. 0;
Matches 4587; Conservative 0; Mismatches 958; Indels 680; Gaps 25;
```

```
QY 958 CCGAGCTCAATATAAAGAGCCCAACCCCTCAGTCGGGGGCGCAGTCTCCGATAGACTG 1017
Db 662 CAGAGCTCAATATAAAGAGCCCAACCCCTCAGTCGGGGGCGCAGTCTCCGATAGACTG 721
QY 1018 CGTCGCGCGGTACCCGTATTCCCAATAAAGCCCTCTTGTCTTTGCAATCCGAATCGTGGA 1077
Db 722 AGTCGCGCGGTACCCGTATTCCCAATAAAGCCCTCTTGTGAG-TTGCATCCGACTGTGCT 780
QY 1078 CTCGCTGATCTTGGGAGGGTCTCCTCAGATTTGATTTGACTGCCCC-ACCTCGGGGGTCTTT 1136
Db 781 CTCGCTGTTCTTGGGAGGGTCTCCTCTGAGTGAATTGACTTACCCGTAGCGGGGTCTTT 840
QY 1137 CATTTGGAGGTTCCACCGAGATTTGGAGACCCCTGCGGAGGACCCGACCCCGCCGCC 1196
Db 841 CATTTGGGGGCTCGTCGGGATCGGAGACCCCTGCGGAGGACCCGACCCGACCCGAC 899
QY 1197 GGGAGGTAAGTCGCCAGCAACTTATCTGTCTGTCCGATGTCGATGTCATGACTG 1256
Db 900 GGGAGGTAAGTCGCCAGCAACTTATCTGTCTGTCCGATTTGTCTGCTATGACTG 959
QY 1257 ATTTTATGCGCTCGCTCGGTACTAGTTAGTAACTAGCTCTGATCTGCGGAGCCCGTG 1316
Db 960 ATTTTATGCGCTCGCTCGGTACTAGTTAGTAACTAGCTCTGATCTGCGGAGCCCGTG 1019
QY 1317 GTGAACTGACGAGTTTCGGAACACCCCGCGCGCAACCTTGGGAGAGCTCCAGGGAATTTCG 1376
Db 1020 GTGAACTGACGAGTTTCGGAACACCCCGCGCGCAACCTTGGGAGAGCTCCAGGGAATTTCG 1079
QY 1377 GGGGCGCTTTTGTGGCGCGCACTGAGTCCCTAAATCCGATCCGATTTAGGACTCTTTGGT 1436
Db 1080 GGGGCGCTTTTGTGGCGCGCACTGAGTCCCTAAATCCGATCCGATTTTGGACTCTTTGGT 1139
QY 1437 GCACCCCTTTAGAGGAGGATATGTGTTCTGTGTAGGAGACGAACTTAAACAGTTC 1496
Db 1140 GCACCCCTTTAGAGGAGGATATGTGTTCTGTGTAGGAGACGAACTTAAACAGTTC 1199
QY 1497 CCGCTCCGCTGTAATTTTGTCTTTTGGGACCCGAAGCCGCGCGCTCTTCTC 1556
```



QY 3713 AGGACACCTACGAGCCCTTACATGACGAGCCCTGCCCTCGCTAACTCGACGGCGCG 3772  
DB 3325 TGCTGCTGCCGACACCACTACTCTGAGCAGCAGCCAGTCCGCGCTGAGCAAGACCCCAACG 3384  
QY 3773 CGGATCCGATAGTCCAAATTTGTTAAAGACAGAGGATATCAGTGGTCCAGGCTCTAGTTT 3832  
DB 3385 AGAGCGCGATCA-----CATGGTCTGCTGAGTTCGTGACCGCGCGCGGAT 3433  
QY 3833 GACTCAACAATATCAACAGCTGAAGCTATAGAGTACGAGCCATAGATAAATAAAGAT 3892  
DB 3434 CACTCTCGCATGGACGAGCTGTACAAGTAAAGCGCGCTCGACGATAAATAAAGAT 3493  
QY 3893 TTTATTTAGTCTCCAGAAAAGGGGGGATGAAGACCCACCTGTAGTTTGGCAAGCT 3952  
DB 3494 TTTATTTAGTCTCCAGAAAAGGGGGGATGAAGACCCACCTGTAGTTTGGCAAGCT 3553  
QY 3953 AGCTTAAGTAAAGCCATTTTCAAGGCATGG-AAAATACATAACTGAGAAATAGAGAAGTT 4011  
DB 3554 AGCTTAAGTAAAGCCATTTTCAAGGCATGG-AAAATACATAACTGAGAAATAGAGAAGTT 3613  
QY 4012 CAGATCAAGGTTAGGAACAGA-GAGACAGCAGAAATATGGGCCAAACAGGATATCTGTGT 4070  
DB 3614 CAGATCAAGGTCAGGAACAGATGAAGTGAATATGGGCCAAACAGGATATCTGTGT 3673  
QY 4071 AAGCGTTCTCTGCCCC-GCTCAGGCCCAAGAACAGTTGGAAACAGAGAAATATGGGCCAAA 4129  
DB 3674 AAGCGTTCTCTGCCCCGCTCAGGCCCAAGAACAGATGGAACAGCTGAATATGGGCCAAA 3733  
QY 4130 CAGGATATCTGTGGTAAAGCAGTTCTAGAGAACCAATCAGATGTTTTCAGGGTGGCCCAA 4249  
DB 3794 GATCGGTCAGCCCTCAGCAGTTCTTAGAGAACCAATCAGATGTTTTCAGGGTGGCCCAA 3853  
QY 4250 GGACTGAAATGACCTGTGCTTATTTGAACCTAACCAATCAGTTCGTTCTCGTCTGTG 4309  
DB 3854 GGACTGAAATGACCTGTGCTTATTTGAACCTAACCAATCAGTTCGTTCTCGTCTGTG 3913  
QY 4310 TTCCGCGCTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCCGCGCGC 4369  
DB 3914 TTCCGCGCTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCCGCGCGCGC 3973  
QY 4370 CAGTCTCCGATAGACTCGTCCCGGGTACCCTGTTCTCAATAAACCCCTCTTGCACT 4429  
DB 3974 CAGTCTCCGATGACTGAGTCCCGGGTACCCTGTTCTCAATAAACCCCTCTTGCACT 4033  
QY 4430 TGCATCCGACTCGTGGTCTCGCTGTTCTTGGGAGGGTCT-CTCTGAGTGATTGACTACC 4488  
DB 4034 TGCATCCGACTCGTGGTCTCGCTGTTCTTGGGAGGGTCTCTCTGAGTGATTGACTACC 4093  
QY 4489 GTCAGCGGGTCTTTCAGTTTCTCCACCTACACAGGCTCTCACTAACATTCCTGATGTG 4548  
DB 4094 GTCAGCGGGTCTTTCAT----- 4113  
QY 4549 CCGCAGGCACTCCGTCAGCCCGTTTTTGTATTAATAAATGCAAGACAGTGTTCCT 4608  
DB 4114 ----- 4113  
QY 4609 TCAAGCCAGACTACATCTGACTCTCGGCTTTTATAAAGAAATGTTGAAGGGCTCTGTGA 4668  
DB 4114 ----- 4113  
QY 4669 CTATCTGCCACAGACTTTTAAAGATTTTATGCTCTCGATGAGGGAATTAAGTCAATC 4728  
DB 4114 ----- 4113  
QY 4729 TATCTCTGCTATTTTGTGGCTTCTCGTATTTTAAATTTCTAGTTTGCCTCCCTTCC 4788  
DB 4114 ----- 4113  
QY 4789 TGAGAGCAGCGGATTGCAGAGTAGTTAATACTCTGAGGGCAGGCTTCTGTGAAAAGGTT 4848

DB 4114 -----TCCGACTTGTGGTCTCGCTGCCTTGGGA 4141  
QY 4849 GCCTGGGCTCAGTGTGAGATTTTGCATAAAAAAGGGGCTCTGCCCTCTGTACACACAGA 4908  
DB 4142 GGGTCTCTCTGAGTGATTGACTACCGCTCAGCGGGGTCTTCAATG----- 4189  
QY 4909 TCGAAATCTAGATGTCATCTACAGATCCCAGGCTTCCGGGGTCTGATCTCAGGGCAT 4968  
DB 4190 ----- 4189  
QY 4969 CTTTGCCTAGAGATCCTCTACGCCGAGCATCGTGGCGGGTACCGAGCTCGAATTCGT 5028  
DB 4190 ----- 4189  
QY 5029 AATCATGTGTATAGTGTCTCTGTGAAATTTGTATCCGCTCACAATTTCCACAACA 5088  
DB 4190 -----CAGCATGTATCAAAATTAATTTGGTTTTTTTCTTAAAGTATTACATTA 4238  
QY 5089 TACGAGCCGGAAGCATAAAGTGTAAAGCTGGGGTGCCTAATGAGTGAAGTAACTCAT 5148  
DB 4239 AA----- 4240  
QY 5149 TAATTTGGTTCGGCTCACTGCCCGCTTTCAGTCCGGAAACCTGTCTGTCGAGCTGCATTT 5208  
DB 4241 -----TGGCCATAGTTGCATTT 4256  
QY 5209 AATGAATCGGCCAACGCGCGGGAGAGCGGTTTTGCGTATTTGGCGCTCTTCCGCTTCT 5268  
DB 4257 AATGAATCGGCCAACGCGCGGGAGAGCGGTTTTGCGTATTTGGCGCTCTTCCGCTTCT 4315  
QY 5269 CGCTCACTGACTCGCTCGCTCGGTTTCGGCTGCGCGAGCGGTATCAGTCACTCAA 5328  
DB 4316 CGCTCACTGACTCGCTCGCTCGGTTTCGGCTGCGCGAGCGGTATCAGTCACTCAA 4375  
QY 5329 AGGCGGTAAATACGTTTATCCACAGATCAGGGGATTAACGAGGAAGAAACATGTGAGCAA 5388  
DB 4376 AGGCGGTAAATACGTTTATCCACAGATCAGGGGATTAACGAGGAAGAAACATGTGAGCAA 4435  
QY 5389 AAGSCCAGCAAAAGCCAGGAACCGTAAAGGCCGGTTCGTGCGTTTTTCCATAGC 5448  
DB 4436 AAGSCCAGCAAAAGCCAGGAACCGTAAAGGCCGGTTCGTGCGTTTTTCCATAGC 4495  
QY 5449 TCCGCCCTCTGACGAGCATCACAAAATCAGACCTCAAGTCAAGTGGGGAACCCGA 5508  
DB 4496 TCCGCCCTCTGACGAGCATCACAAAATCAGACCTCAAGTCAAGTGGGGAACCCGA 4555  
QY 5509 CAGGACTATAAGATACAGCGGTTTTCCCTCGGAAAGCTCCCTCGTGGGCTCTCTCTGTT 5568  
DB 4556 CAGGACTATAAGATACAGCGGTTTTCCCTCGGAAAGCTCCCTCGTGGGCTCTCTCTGTT 4615  
QY 5569 CGACCTCGCGCTTACCGGATACCTGTCGCTTTCTCCCTTCGGGAGCGTGGCGTTTT 5628  
DB 4616 CGACCTCGCGCTTACCGGATACCTGTCGCTTTCTCCCTTCGGGAGCGTGGCGTTTT 4675  
QY 5629 CTCTAGCTCAGCTGTAGGTATCTCAGTTCGCTGTAGTTCGTTCCGCTCAAGTGGGCT 5688  
DB 4676 CTCTAGCTCAGCTGTAGGTATCTCAGTTCGCTGTAGTTCGTTCCGCTCAAGTGGGCT 4735  
QY 5689 GTGTGCAAGAACCCCGCTTACGCCGACCGCTCGGCTTATCCGGTAACTATCTGTTG 5748  
DB 4736 GTGTGCAAGAACCCCGCTTACGCCGACCGCTCGGCTTATCCGGTAACTATCTGTTG 4795  
QY 5749 AGTCAACCCCGTAAAGACAGACTTATCGCCACTGGGAGCAGCCACTGGTAAACAGGATTA 5808  
DB 4796 AGTCAACCCCGTAAAGACAGACTTATCGCCACTGGGAGCAGCCACTGGTAAACAGGATTA 4855  
QY 5809 GCAGAGCAGGTATGTAGGCGGTCTACAGAGTTCTTGAAGTGGTGGGCTTAACCTACGCT 5868  
DB 4856 GCAGAGCAGGTATGTAGGCGGTCTACAGAGTTCTTGAAGTGGTGGGCTTAACCTACGCT 4915  
QY 5869 AACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAA 5928

4916 A C A C T A G A A G G C A G A T T T T G G T A T C T G C G C T C T G T A A G C C A G T T A C C T T C G G A A A A 4975  
QY  
5929 G A G T T G G T A G C T T C A T C C G C A A A C A A C C A C C G C T G T A G C G T G T T T T T T G T T T 5988  
Db  
4976 G A G T T G G T A G C T T C T G A T C C G G A A A C A A C C A C C G C T G T A G C G T G T T T T T G T T T 5035  
QY  
5989 G C A A G C A G C A G A T T A C G C G A G A A A A A A A G G A T C T C A G A A G A T C C T T T G A T C T T T C T A 6048  
Db  
5036 G C A A G C A G C A G A T T A C G C G A G A A A A A A A G G A T C T C A G A A G A T C C T T T G A T C T T T C T A 5095  
QY  
6049 C G G G G T C T G A C G C T C A G T G G A A C A A A A C T C A G T T A A G G A T T T T G T C A T G A G A T A T 6108  
Db  
5096 C G G G G T C T G A C G C T C A G T G G A A C A A A A C T C A G T T A A G G A T T T T G T C A T G A G A T A T 5155  
QY  
6109 C A A A A G G A T C T C A C C T A G A T C C T T T T A A A T T A A A A T G A G T T T T T T T T T T T T T T T T 6165  
Db  
5156 C A A A A G G A T C T T C A C C T A G A T C C T T T T A A A T T A A A A A T G A A G T T T G C G C A A T C A A T C T 5215  
QY  
6166 A A A G T A T A T A G T A A A A C T T G G T C T G A C A G T T A C C A A T G C T T A A T C A G T G A G G C A C T A 6225  
Db  
5216 A A A G T A T A T A G T A A A C T T G G T C T G A C A G T T A C C A A T G C T T A A T C A G T G A G G C A C T A 5275  
QY  
6226 T C T A C G C A T C T G T C T A T T T C G T T C A T C C A T A G T T G C T G A C T C C C C G T C G T G T A G A T A A 6285  
Db  
5276 T C T A C G C A T C T G T C T A T T T C G T T C A T C C A T A G T T G C T G A C T C C C C G T C G T G T A G A T A A 5335  
QY  
6286 C T A C G A T A G G A G G C T T A C C A T C T G G C C C A G T G C T C A A T G A T A C C G C A G A C C C A C 6345  
Db  
5336 C T A C G A T A G G A G G C T T A C C A T C T G G C C C A G T G C T G C A A T G A T A C C G C A G A C C C A C 5395  
QY  
6346 G C T C A C C G C T C C A G A T T A T C A G C A A T A A C A G C C A C C G A A G G C C G A G C C G A G A A 6405  
Db  
5396 G C T C A C C G C T C C A G A T T A T C A G C A A T A A C A G C C A C C G A A G G C C G A G C C G A G A A 5455  
QY  
6406 G T G G T C C T C A A C T T A T C G C C T C C A T C C A G T C T A T T A A T T G T G C G G G A A G C T A G A G 6465  
Db  
5456 G T G G T C C T C A A C T T A T C G C C T C C A T C C A G T C T A T T A A T T G T G C G G G A A G C T A G A G 5515  
QY  
6466 T A A G T A G T T C G C A G T T A T A G T T T G C G A A A G T T T G C C A T T G C T A C A G G C - T G T G G 6524  
Db  
5516 T A A G T A G T T C G C A G T T A T A G T T T G C G A A A G T T T G C C A T T G C T A C A G G C A T C G T G G 5575  
QY  
6525 T G T C A G C T C G T G T T G T A T G G C T T C A T T C A G C T C C G T T C C C A A C G A T C A A G C G A G 6584  
Db  
5576 T G T C A G C T C G T G T T G T A T G G C T T C A T T C A G C T C C G T T C C C A A C G A T C A A G C G A G 5635  
QY  
6585 T T A C A T G A T C C C C A T T T G T G C A A A A A A G C G T T A G C T C C T T C G G T C C T C C G A T C G T T G 6644  
Db  
5636 T T A C A T G A T C C C C A T T T G T G C A A A A A A G C G T T A G C T C C T T C G G T C C T C C G A T C G T T G 5695  
QY  
6645 T C A A A G T A A G T T G C C G A G T T A T C A T C A T A G T T A T G G C A G C A C T G C A T A A T T C T C 6704  
Db  
5696 T C A A A G T A A G T T G C C G A G T T A T C A T C A T G T T A T G C A G C A C T G C A T A A T T C T C 5755  
QY  
6705 T T A C T G C A T G C C A T C C G T A A G A T G C T T T T C T G A C T G T G A C T C A A C C A G T C A T 6764  
Db  
5756 T T A C T G C A T G C C A T C C G T A A G A T G C T T T T C T G A C T G T G A C T A C A A C A A G T C A T 5815  
QY  
6765 T C T G A A T A G T A T G C G C A C C G A G T T G C T T T T G C C G C G T C A A T A C C G G A T A A T A 6824  
Db  
5816 T C T G A A T A G T A T G C G C A C C G A G T T G C T T T G C C G C G T C A A C C G G A T A A T A 5875  
QY  
6825 C G C G C C A C A T A G C A A C T T T A A A A G T G C T A T C A T T T G A A A A A C G T T C T T C G G G C G A A 6884  
Db  
5876 C G C G C C A C A T A G C A A C T T T A A A A G T G C T A T C A T T T G A A A A C G T T C T T C G G G C G A A 5935  
QY  
6885 A A C T C T C A G A G T T T A C C G T T G A G A T C C A G T T C G A T G A A C C C A C T C G T G C A C C C A 6944  
Db  
5936 A A C T C T C A G A G T T T A C C G T T G A G A T C C A G T T C G A T G A A C C C A C T C G T G C A C C C A 5995  
QY  
6945 A C T G A T C T T C A G C A T C T T T A C T T T C A C C A G G G T T T C T G G T G A G A A A A C A G G A A G C 7004  
Db  
5996 A C T G A T C T T C A G C A T C T T T A C T T T C A C C A G G G T T T C T G G T G A G A A A A C A G G A A G C 6055

QY 7005 A A A A T G C G C A A A A A A A A G G A T A A G G C G A C A C G A A A T G T T G A A T A C T A T C T A C T T C C 7064  
Db 6056 A A A A T G C G C A A A A A A A A G G A T A A G G C G A C A C G A A A T G T T G A A T A C T A T C T A C T T C C 6115  
QY 7065 T T T T T C A A T A T T A T T A A G C A T T T A T C A G G G T T A T T G T C T C A T A G C G A T A C A T A T T T G 7124  
Db 6116 T T T T T C A A T A T T A T T A A G C A T T T A T C A - G G T T A T T G T C T A T A G C G A T A C A T A T T T G 6174  
QY 7125 A A T G T A T T A G A A A A T A A A C A A A T A G G G T T C C G C G A C A T T T C 7169  
Db 6175 A A T G T A T T A G A A A A T A A C A A A T A G G G T T C C G C G A C A T T T C 6219

RESULT 14  
US-09-826-025-1  
; Sequence 1, Application US/09826025  
; Patent No. US20020162123A1  
; GENERAL INFORMATION:  
; APPLICANT: Chang, Lung-Ji  
; TITLE OF INVENTION: Combination Immunogene Therapy  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/826,025  
; FILING DATE: 04-Apr-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/838,702  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ingolia, Diane E.  
; REGISTRATION NUMBER: 40,027  
; REFERENCE/DOCKET NUMBER: CHANG-02687  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6145 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-826-025-1

Query Match 40.0%; Score 3062.2; DB 9; Length 6145;  
Best Local Similarity 71.7%; Pred. No. 0;  
Matches 4818; Conservative 0; Mismatches 1168; Indels 735; Gaps 32;  
QY 549 TTTTGAAGAGACCCACCTGTAGGTTTGGCAAGCTAGCTTAAGTAACGCCATTTCGCAAGG 608  
Db 144 TTTTGAAGAGACCCACCCCGTAGG--TGGCAAGCTAGCTTAAGTAACGCCACTTTTGAAGG 201  
QY 609 CATGG-AAAATACATACTGAGATAGAGAGCTTCAGATCAAGGTTAGGAACAGAGAGAC 667  
Db 202 CATGAAAATAATACATACTGAGATAGAAAAGTTCAGATCAAGGTCAGGAACAAAGAAC 261  
QY 668 AGCAGATATGGGCCAAACAGGATATCTGTGTAAAGAGCTTCCTGCCCC-GTTCAGGGCC 726  
Db 262 AGCTGAATA---CCAAACAGGATATCTGTGTAGCGTTCTCTGCCCGGCTCAGGGCC 317







Db 2356 GCGAATGGGCTGACCGCTCTCTCGTCTTACGGTATCGCGCTCCC----- 2402  
 QY 2942 CATATTGGATGAGTTGGGTGAGACAGCAGCACTGGAAAAAGTCTTGAGTGGATTGGAGAA 3001  
 Db 2403 -----GATTCGACGCGCATCGCCCTTCTATCGCTTCTTGACGAG 2441  
 QY 3002 TTCACTCAGATAGCAGTACGATTAACTATGCGCGCTCTTAAAGGATAGATTACAAATAT 3061  
 Db 2442 TTCTTCTGAGCGGACTCTGGGGTTCGAAATGACCGACCAAGCAGCCCAACCTGCCAT 2501  
 QY 3062 CCGCAGAACGCGCAAGAACACATTGTTCTGCAAAATGGACAGCCTTGAGACCCGGAAGACA 3121  
 Db 2502 CACGAGATTTCGATTCCACGCGCGCTCTATGAAGGTTGGCTTCGGAATCGTTTTCC 2561  
 QY 3122 CCGGGTCTATTTTGTGCAAGCTTTACTTGGCTTCCCTGGTTTGTCTTATGAGGCC 3181  
 Db 2562 GCGACGCC-----GGCTGGATGATCCT 2583  
 QY 3182 AAGGACCCCGGTCACCGTCTCCAGTCTAAGCCACACAGCGCCAGCGCGCGACAC 3241  
 Db 2584 CAGCGCGGGATCTCATCTGAGTCTTTCGCCACCCCGGCTCGATCCCTCGCGAG 2643  
 QY 3242 CAACACCGCGGCCACCATCGGTCGAGCGCTGTCAGCCCTGTCCTGCGCCAGAGCGGCTCGGC 3301  
 Db 2644 TTGGTTCACTGCTGCGCTGAGGCTGACGACCTCG----- 2678  
 QY 3302 CAGCGCGGGGCGCAGTGACACAGAGGGGTGACCTTTCGCTGGATCCCAACTCT 3361  
 Db 2679 -----CGGAGTTCTACCGGCAGTGCAAAATCG 2705  
 QY 3362 GCTACCTGCTGATGAATCTCTTCTCATCTATGGTGTCTATTCTCATGCGCTTGTCTGA 3421  
 Db 2706 TCGCATCCAGGA---AACACAGCGGCTATCCGCGCATCTCGCGCCCGAACTGCAG 2762  
 QY 3422 GAGTGAAGTTTCAGCAGGAGCGAGAGCCCGCGGTACAGCAGAGGCGCAGAACCCAGCTCT 3481  
 Db 2763 AGTGGGAGCAGATGCGCGCTTGGTGCACCGGAGCGCTCTCTGCGCTGATAC 2822  
 QY 3482 ATAAACAGCTCAATCTAGACGAAGAGAGAGTACGATGTTTGGACGAAGAGAGTGGCC 3541  
 Db 2823 AGAACGAATGCTTGACGGCATCTCATGATGTGTCTT-----CCCGTTTTT 2869  
 QY 3542 GGGACCTCAGATGGGGGAAAGCGAGAGGAGAACCCCTCAGGAGGCGCTGTACAATG 3601  
 Db 2870 CGCTGAGTCACTCGTGGATGAGCGCTGGCGCTGCTGCGGACGCGAG-----CTG 2925  
 QY 3602 AACTGCAGAAAGATAAGATGGGGAGCGCTACAGTGAGATGGGATGAAAGCGAGCGCC 3661  
 Db 2926 CTCACCACTCGAGGCGGTGACGCGCTGCAGAGCGCGAGTGCAGAACTGC----- 2978  
 QY 3662 GGAGGGGAGGGGACGATGGCTTTACAGGCTCTAGTACAGCCACCAAGGACACT 3721  
 Db 2979 -----TCCAAAGGAGCTCAAGGCTTTCCAGGAGCAGCTAGGCTGACTCCATCGAGCCAG 3033  
 QY 3722 ACGAGCCCTTCACTGACGAGGCGCTGCCCCCTCGCTAACTCGACGCGCGCGATCCGG 3781  
 Db 3034 TGTA-----GAGATAAGCTTATCG 3052  
 QY 3782 ATTAGTCCCAATTTGTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTTGACTCAACA 3841  
 Db 3053 ATTAGTCCCAATTTGTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTTGACTCAACA 3112  
 QY 3842 ATATCAACGAGTGAAGCTTATAGATACGAGCCATAGATAAAATAAAGATTTATTTAG 3901  
 Db 3113 ATATCAACGAGTGAAGCTTATAGATACGAGCCATAGATAAAATAAAGATTTATTTAG 3172  
 QY 3902 TCTCCAGAAAGGGGGGAAAGAGCCACCTGATAGTTTGGCAGCTAGCTTAAGT 3961  
 Db 3173 TCTCCAGAAAGGGGGGAAAGAGCCACCTGATAGTTTGGCAGCTAGCTTAAGT 3232  
 QY 3962 AACGCCATTTGCAAGGCATGG-AAAAATACATAACTGAGAAATAGAGAAGTTCCAGATCAAG 4020

Db 3233 AACGCCATTTTGAAGGCGATGGAAAAATACATAACTGAGAAATAGAGAAGTTCCAGATCAAG 3292  
 QY 4021 GTTAGGAACAG- GAGACAGCAGATATGGGCAACACAGGATATCTGTGGTAAAGCAGTTTC 4079  
 Db 3293 GTCAGGAACAGATGGAACAGCTGAATATGGGCAACACAGGATATCTGTGGTAAAGCAGTTTC 3352  
 QY 4080 CTGCCCC--GCTCAGGCGCAAGAACAGTTGGAAACAGGAAATATGGGCCAAACAGGATATC 4138  
 Db 3353 CTGCCCCGGCTCAGGCGCCAGAACAGATGGAACACAGCTGAATATGGGCCAAACAGGATATC 3412  
 QY 4139 TGTGTAAGCAGTTCTCTGCCCGCTCAGGCGCAAGAACAGATGTTCCAGATCGGTC 4198  
 Db 3413 TGTGTAAGCAGTTCTCTGCCCGCTCAGGCGCAAGAACAGATGTTCCAGATCGGTC 3472  
 QY 4199 CCGCCCTCAGCAGTTTCTAGAGAACATCAGATGTTTCCAGGCTGCCCAAGCACTGAA 4258  
 Db 3473 CAGCCCTCAGCAGTTTCTAGAGAACATCAGATGTTTCCAGGCTGCCCAAGCACTGAA 3532  
 QY 4259 ATGACCCCTGTGCTTATTTGAACTAAACCAATCAGTTTCTGCTTCTGTTTCCGGCGC 4318  
 Db 3533 ATGACCCCTGTGCTTATTTGAACTAAACCAATCAGTTTCTGCTTCTGTTTCCGGCGC 3592  
 QY 4319 TTCTGCTCCCCGAGCTCAATAAAGAGCCCAACACCCCTCACTCGCGCGCCAGCTCTCC 4378  
 Db 3593 TTCTGCTCCCCGAGCTCAATAAAGAGCCCAACACCCCTCACTCGCGGCGCCAGTCTCC 3652  
 QY 4379 GATAGACTGCTGCGCGCGGTACCCGTTTCTCAATAAACCCCTCTTGCAGTTGTCATCCGA 4438  
 Db 3653 GATTGACTGAGTGCAGCGGATACCCGTTGATCAATAAACCCTCTTGCAGTTGTCATCCGA 3712  
 QY 4439 CTCGTGCTGCTGCTGCTTCTTGGGAGGCTCT--CTCTGAGTGAATGACTTACCCCTCAGCGG 4497  
 Db 3713 CTGTGCTGCTGCTGCTTCTTGGGAGGCTCTCTCTGAGTGAATGACTTACCCCTCAGCGG 3772  
 QY 4498 GGTCTTTCAGTTTCTCCACCTTACACAGCTCTCACTAAACATCTCTGATGTGCGCGCAGGA 4557  
 Db 3773 GGGTCTTTCATTTGGGGGCTCGTCCGGATCGGAGACCCCTGCGCAGGACCAACCGACC 3832  
 QY 4558 CTCGCTCAGCCCGGTTTTTGTATTATAAATAAATGCAAGAACAGTGTTCCTTCAAGCCAG 4617  
 Db 3833 CACCACCGGAGGTAAAGCTG----- 3852  
 QY 4618 ACTACATCTGACTCTCGGCTTTATAAAGAAATGTTGAAGGCTCTGTGGAATCTATCGC 4677  
 Db 3853 -----GCTGCTCGCGGTTTTCGGTGATGACGTGAAACCTCTGAC-ACATGCACT 3904  
 QY 4678 ACAGCACTTTTAAAGATTTTATGCTCTCTGATGAGGATTTAGTCAATCTATCTCTCT 4737  
 Db 3905 CCGGAGACGGTCAACAGCTTGTCTGTAAGCGATGCGCGGAGCAGACAAGCCCGTCAGGG 3964  
 QY 4738 CTATTTTGTGCTGCTCTCGTATTTTAAATTTTCTAGTTTGCATCTCTTCTGAGAGCAC 4797  
 Db 3965 CGCGTCAGCGGTTGTGGG----- 3984  
 QY 4798 GSCGATGACAGATGATTAAATCTCTGAGGCGAGGCTTCTGTGAAAAGTTGCTGCGGCT 4857  
 Db 3985 ----- 3984  
 QY 4858 CAGTGTGAGATTTTCCATATAAAGGGGTCTGCGCCCTGTGTACAGACAGATCGGAATCT 4917  
 Db 3985 ----- 3984  
 QY 4918 AGAGTGCATATCAGAGTCCCGCGGTTCCGGGCTCTGATCTCAGGGCATCTTTGCTTA 4977  
 Db 3985 -----GGTCTCGGGGCGCAGCCATGACCCAGTC----- 4012  
 QY 4978 GAGATCTCTAGCCCGGACGATCTGCGGTACCGGTACCGAGCTCGAATTCGTAATCATGCT 5037  
 Db 4013 -----ACGTAGCAGTACGAGTGTATCTGCTTAACTATGCGGCT 4055  
 QY 5038 CATAGCTCTTCTGTGTGAAATTTGTTATCCGCTCACAATTTCCACACATACGAGCGG 5097  
 Db 4056 CAGACAGATTTGACTGAGAG----- 4076

Qy	5098	GAAGCATAAGTGTAAGAACCGCTGGGGTGCCTAAATGAGTGAGCTAACTCACATTAATTGCGT	5157
Db	4077	-----	4076
Qy	5158	TGGCTCACTGCCCGCTTTCCAGTCGCGGAACCTGTGTCGCCAGCTGCATTAATGATCG	5217
Db	4077	-----TGCACTATATCGGTGTGAATAC	4100
Qy	5218	GCCAAACGCGGGGAGAGGGGTTTGCGTATTTGGGCGCTCTTCGCTTCCTCGCTCACTG	5277
Db	4101	CGCACAGATCGTAAGGAGAAAATACCGCATCAGGCGCTCTTCGCTTCCTCGCTCACTG	4160
Qy	5278	ACTCGCTGCGTTCGGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAAGCGGTAA	5337
Db	4161	ACTCGCTGCGTTCGGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAAGCGGTAA	4220
Qy	5338	TACGGTTATCCACAGATCAGGGGATAACGACAGGAAAGACATGTGAGCAAAAGGCCAGC	5397
Db	4221	TACGGTTATCCACAGATCAGGGGATAACGACAGGAAAGACATGTGAGCAAAAGGCCAGC	4280
Qy	5398	AAAAGGCCAGGAACCGTAAAAAGCCGCTTGCTGGCGTTTTCATAGGCTCCGCCCCC	5457
Db	4281	AAAAGGCCAGGAACCGTAAAAAGCCGCTTGCTGGCGTTTTCATAGGCTCCGCCCCC	4340
Qy	5458	CTGACGAGCATCAAAAATCGAGCTCAAATCAGAGTGGCGAAACCCGACAGGACTAT	5517
Db	4341	CTGACGAGCATCAAAAATCGAGCTCAAATCAGAGTGGCGAAACCCGACAGGACTAT	4400
Qy	5518	AAAGATACAGGCGTTTCCCTTGGAAGCTCCCTCGTGGCGCTCTCCTGTTCCGACCCCTGC	5577
Db	4401	AAAGATACAGGCGTTTCCCTTGGAAGCTCCCTCGTGGCGCTCTCCTGTTCCGACCCCTGC	4460
Qy	5578	CGCTTACCGGATACCTGTCCGCCCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCT	5637
Db	4461	CGCTTACCGGATACCTGTCCGCCCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCT	4520
Qy	5638	CAGCTGTAGGTATCTCAGTTTCGGTGTAGGTTCGTTTCGCTCAAAGCTGGGCTGTGTGCACG	5697
Db	4521	CAGCTGTAGGTATCTCAGTTTCGGTGTAGGTTCGTTTCGCTCAAAGCTGGGCTGTGTGCACG	4580
Qy	5698	AACCCCGCTTCAGCCCGACCGCTGCGCCTTATCCCGGTAACTATCGTCTTTAGTCCCAACC	5757
Db	4581	AACCCCGCTTCAGCCCGACCGCTGCGCCTTATCCCGGTAACTATCGTCTTTAGTCCCAACC	4640
Qy	5758	CGGTAAACACGACTTATCGCCACTGGCAGCAGCCACTGTGTAACAGGATTTAGCAGAGCGA	5817
Db	4641	CGGTAAACACGACTTATCGCCACTGGCAGCAGCCACTGTGTAACAGGATTTAGCAGAGCGA	4700
Qy	5818	GGTATGTAGGGGTCTACAGATTCTTGAAGTGGTGGCCCTTACTACGGCTACACTAGAA	5877
Db	4701	GGTATGTAGGGGTCTACAGATTCTTGAAGTGGTGGCCCTTACTACGGCTACACTAGAA	4760
Qy	5878	GGACAGTATTTGGTATCTCGGCTCTGCTGAAGCCAGTTTACTCTTCGGAATAAGAGTTGGTA	5937
Db	4761	GGACAGTATTTGGTATCTCGGCTCTGCTGAAGCCAGTTTACTCTTCGGAATAAGAGTTGGTA	4820
Qy	5938	GCTCTTGATCCGGCAAAACAAACCAACCGCTGGTAGCGGTGGTTTTTTTGTTCGAAAGCAGC	5997
Db	4821	GCTCTTGATCCGGCAAAACAAACCAACCGCTGGTAGCGGTGGTTTTTTTGTTCGAAAGCAGC	4880
Qy	5998	AGATTACCGCAGNAAAAAAGGATCTCAAGAGATCCTTTGATCTTTTCTACCGGGTCTG	6057
Db	4881	AGATTACCGCAGNAAAAAAGGATCTCAAGAGATCCTTTGATCTTTTCTACCGGGTCTG	4940
Qy	6058	ACGCTCAGTGGACGAAAACTCAGTTTAAAGGATTTTGGTTCATGAGATTATCAAAAGGA	6117
Db	4941	ACGCTCAGTGGACGAAAACTCAGTTTAAAGGATTTTGGTTCATGAGATTATCAAAAGGA	5000
Qy	6118	TCCTCACTAGATCCTTTTAAATTTAAAAATGAAGTTTTTAAATCAATCAAGTATATATG	6177
Db	5001	TCCTCACTAGATCCTTTTAAATTTAAAAATGAAGTTTTTAAATCAATCAAGTATATATG	5060

Qy	6178	AGTAAACTTGGCTGTGACAGTTTAAACAATGCTTTAATACGTGTAGGCACTTATCTCAGCGATCT	62337
Db	5061	AGTAAACTTGGCTGTGACAGTTTAAACAATGCTTTAATACGTGTAGGCACTTATCTCAGCGATCT	5120
Qy	6238	GTCTATTTCCTGTCATCCATAGTTGGCTGACTCCCGTCTGTTAGATAAATACCATACGG	6297
Db	5121	GTCTATTTCCTGTCATCCATAGTTGGCTGACTCCCGTCTGTTAGATAAATACCATACGG	5180
Qy	6298	AGGCTTTACCATTCTGGCCCCAGTGTGCTGCAATGATACCGGAGACCCACGCTCACCGGCTC	6357
Db	5181	AGGCTTTACCATTCTGGCCCCAGTGTGCTGCAATGATACCGGAGACCCACGCTCACCGGCTC	5240
Qy	6358	CAGATTATCAGCAATAAACACGACGAGCGGAAAGGCGCGAGCGGAGAGTGTCTTGCAG	6417
Db	5241	CAGATTATCAGCAATAAACACGACGAGCGGAAAGGCGCGAGCGGAGAGTGTCTTGCAG	5300
Qy	6418	CTTTATCCGGCTCCATCCAGTCTATTAAATTTGTTGCGGGAAGCTACGAGTAAGTAGTTCG	6477
Db	5301	CTTTATCCGGCTCCATCCAGTCTATTAAATTTGTTGCGGGAAGCTACGAGTAAGTAGTTCG	5360
Qy	6478	CAGTTAATAGTTTGGCGCAAGTGTGTGCCATTGCTACAGGC-TCGTGGTGTCTCAGCTCGT	6536
Db	5361	CAGTTAATAGTTTGGCGCAAGTGTGTGCCATTGCTACAGGCATCGTGGTGTCTCAGCTCGT	5420
Qy	6537	CGTTTGTATGCTTCAATTCAGTCTCCGGTTCCTCCAAAGATCAAGCGAGTTACATGATCCC	6596
Db	5421	CGTTTGTATGCTTCAATTCAGTCTCCGGTTCCTCCAAAGATCAAGCGAGTTACATGATCCC	5480
Qy	6597	CCATGTTGTCAAAAAGCGGTTAGCTCTTCCGTTCCGATCCGTTGTCAGAGTAAGT	6656
Db	5481	CCATGTTGTCAAAAAGCGGTTAGCTCTTCCGTTCCGATCCGTTGTCAGAGTAAGT	5540
Qy	6657	TGGCCGAGTGTATCACTCATGCTTATGGCAGCACTGCAATAATCTCTTACTGTCTATGC	6716
Db	5541	TGGCCGAGTGTATCACTCATGCTTATGGCAGCACTGCAATAATCTCTTACTGTCTATGC	5600
Qy	6717	CATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACAAAGTCAATCTGAGAATAGT	6776
Db	5601	CATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACAAAGTCAATCTGAGAATAGT	5660
Qy	6777	GTATCGGCGCACCGAGTTGCTTTGCGCGGCTCAATACGGGTAATACCGCGCCACATA	6836
Db	5661	GTATCGGCGCACCGAGTTGCTTTGCGCGGCTCAATACGGGTAATACCGCGCCACATA	5720
Qy	6837	GCAGACTTTAAAGTGCTCATCATTTGGAAAAAGTTCTTCGGGCGCAAAACTCTCAAGGA	6896
Db	5721	GCAGACTTTAAAGTGCTCATCATTTGGAAAAAGTTCTTCGGGCGCAAAACTCTCAAGGA	5780
Qy	6897	TCATTACCGCTGTGAGATCCAGTTCGATGPAACCCACTCGTGCACCCAACTGATCTTCAG	6956
Db	5781	TCATTACCGCTGTGAGATCCAGTTCGATGPAACCCACTCGTGCACCCAACTGATCTTCAG	5840
Qy	6957	CATCTTTTACTTTTCCACGAGGTTTCTGGGTGAGCAAAAACAGGAGGCAAAATGCCGCA	7016
Db	5841	CATCTTTTACTTTTCCACGAGGTTTCTGGGTGAGCAAAAACAGGAGGCAAAATGCCGCA	5900
Qy	7017	AAAAGGAATAAGGGGACACGGAATGTTGAATACTCATCTCTCTCTCTCTCTCTCTCTCTCT	7076
Db	5901	AAAAGGAATAAGGGGACACGGAATGTTGAATACTCATCTCTCTCTCTCTCTCTCTCTCTCT	5960
Qy	7077	ATTGAAGCAATTTATCAGGGTTATTGTCTCATGACGGGATACATATTTGAATGTATTTAGA	7136
Db	5961	ATTGAAGCAATTTATCAGGGTTATTGTCTCATGACGGGATACATATTTGAATGTATTTAGA	6020
Qy	7137	AAAATAACAAATAGGGGTTCCGCGCACATTTCCCGGAAAAGTGCCACCTGACGCTTAAG	7196
Db	6021	AAAATAACAAATAGGGGTTCCGCGCACATTTCCCGGAAAAGTGCCACCTGACGCTTAAG	6080
Qy	7197	AAACCATTTATTATCATGACATTAACCTTATAAAATAGGCGTATCACAGGCCCTTTTCGTC	7256
Db	6081	AAACCATTTATTATCATGACATTAACCTTATAAAATAGGCGTATCACAGGCCCTTTTCGTC	6140
Qy	7257	T 7257	

7257 T 7257



Db 1492 TGTACACCTAAGCCTCGCCCTCTCTCTCTCCATCCGCCCGGTCTCTCCCCCTTGAACC 1551  
Qy 1982 TCCTGTTGCGACCCCGCTCGATCCTCCCTTTATCCAGCCCTCACTCTTCTCTAGGCGC 2041  
Db 1552 TCCTGTTGCGACCCCGCTCGATCCTCCCTTTATCCAGCCCTCACTCTTCTCTAGGCGC 1611  
Qy 2042 CCCCATATGGCCATATGAGATCTATATGGGGCACCCCGCCCTTGTAAACTTCCCTGA 2101  
Db 1612 C-----GGAAATTCGATCTGA 1627  
Qy 2102 CCCTGACATGACAAGAGTTACTTAACAGACCCCTCTCTCAAGCTCACTTACAGGCTTCTAC 2161  
Db 1628 TCAAGAGACAGGATGAGATGCTTTCGATGATGATGAACAAGATGATGACGAGGTTTC 1687  
Qy 2162 TTAGTCCAGCAGCAAGTCTGGAGACTCTGGCGGAGCTTACCAAGAAACAATGACCGGA 2221  
Db 1688 TCCGCGCGCTGG-----GTGGAGAGCTATTTCGGCTATGACTGGGCAAA 1733  
Qy 2222 CGGTGGTACCTCAACCTTACCGAGTCGGCGACACAGTGTGGGTCCGCGGACACAGACT 2281  
Db 1734 CAGACAATCGGCTGCTCTGATCCCGCTGTTCGGCTGTGAGCGCAGGGCGGCCGGTT 1793  
Qy 2282 AAGAACCTAGAACCTCGTGGGAAGGACCTTACACAGTCTGCTGACACACCCCGCC 2341  
Db 1794 CTTTGTGTCAGACCGACCTGTGCGGTGCTGCTGAATGAAGTGC----- 1836  
Qy 2342 CTCAAAGTAGACGGCATCCGAGCTGGATACACGCGCGCCACGCTGAAGGCTGCCGACCCC 2401  
Db 1837 -----AGACGAGGACGCGGCTATCGTGGCTGGCCACGACGCGGCTTCTTGGCAGC 1891  
Qy 2402 GGGGTGACCATCCTCTAGACTGCTCATGCGATGGAGTGTATCATCTCTTCTTGGTAG 2461  
Db 1892 TGTGCTGACGTTGTCACTGAAGCGGAAGGACCTGGCTGTATTTGGCGGAAGTGC CGGG 1951  
Qy 2462 CAACAGCTACAGTGTCTCACTCCGACATCCAGCTGACCCAGAGCCAGCAGCTGACGG 2521  
Db 1952 GCAGGATCTCTGTCTCATCTCACTTGTCTCTGCGAGAAAGTATCCATCATGCTGATGC 2011  
Qy 2522 CCAGCGTGGGTGACAGAGTGACCATCACTGTGAAGCGCAGTCAGGATGTGGTACTTCTG 2581  
Db 2012 AATGCGCGGCTGCATAGCTGTGATCCGGCTACCTGCGCCATTCGACACCAAGCGAACA 2071  
Qy 2582 TAGCTTGTGTAACAGCAGAGCAGGTAAAGGTCCAAAGCTCTGATCTACTTGAATCCA 2641  
Db 2072 TCGCATCAGCGCAGCAGCTACTCGGATGGAAGCGGCTCTTGTGATCAGGATGATCTGA 2131  
Qy 2642 CCGGCACTGTGTGTCAGACAGATTCAGGGGTAGCGGTAGCGGTACCGACTTCACCT 2701  
Db 2132 CGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGT-----TCGCCAGGC 2175  
Qy 2702 TCACCATCAGCAGCCTCCAGCCAGAGGACATCGCCACCTACTACTGCGCAGCAATATAGCC 2761  
Db 2176 TCAAGCGCGCATGCCCAGCGGAGGATCTCGTGTGACCAATGCGGATCGCTGCTTGC 2235  
Qy 2762 TCTATCGGTCTGTGCGCAAGGACCAAGGTGGAATCAACAGAGGTGGCTCAGGATCGG 2821  
Db 2236 CGAATATCATGTGGAATAATGSCCGCTTTCTGGATTCTATCGACTGTGCGCGCTGGGTG 2295  
Qy 2822 TGTGATCGGCTCTGGTGGCTCAGATCGGAGGTCCAACTGGTGAGAGCGGTGAGGTG 2881  
Db 2296 TGGCGGACCGCTATCAGGACATAGGTTGGCTTACCCGCTGATATTGCTGGAAGACTTGGCG 2355  
Qy 2882 TTGTGCACTGGCGGCTCCCTGCGCTGCTCTGCTCGCATCTGGCTTCGATTTCAACA 2941  
Db 2356 GCGAATGGGCTGACCGCTTCTCTGTGCTTTACGGTATCGCGCTCC----- 2402  
Qy 2942 CATATTGATGATTTGGGTGAGACAGGCACCTGGAAAAGTCTTGTAGTGGATTGGAGAAA 3001  
Db 2403 -----GATTCGACGGCATCGCTTCTATCGCTTCTTGAACGAG 2441  
Qy 3002 TTCTATCCAGATGACGATGATTAATCTATGCGCGCTCTCTAAAGATAGATTTCAATAT 3061  
Db 2442 TTCTTCTGAGCGGGACTCTGGGGTTGGAATGACCGACCAAGCGACGCGCCAACTGCCAT 2501

Qy 3062 CGCAGAGCAACCGCAAGAACATTTGTTCTTGCAAAATGGAAGAGCTGAGACCCCGAAGACA 3121  
Db 2502 CACGAGATTTGATTCACACCGCCGCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTCC 2561  
Qy 3122 CCGGGGTCTATTTTGTGCAAGCCTTTACTTCGGCTTCCCTGGTTTGTCTATTGGGGCC 3181  
Db 2562 GGSAGGCC-----GGCTGGATGATCCT 2583  
Qy 3182 AAGGGACCCCGGTCTCCAGTCTCCAGTCTTACGCCACCAAGCGCGCAGCGCGGACAC 3241  
Db 2584 CCAGCGGGGGATCTCATGCTGGAGTTCTTCGCCCAACCCCGGGCTCGATCCCTCGCGAG 2643  
Qy 3242 CAACACCGCGGCCCAACCATCGCGTTCGAGCCCTTGTCTCCGCGCCAGAGCGGCTCGGC 3301  
Db 2644 TTGGTTCAAGCTGCTGCTGAGGCTGGACGACTCG----- 2678  
Qy 3302 CAGCGCGGGGGCGCAGTGCACAGAGGGGGCTGGACTTCGCCCTGATCCCAAACTCT 3361  
Db 2679 -----CGAGTTCTACCGGCAGTGCAAATCCG 2705  
Qy 3362 GCTACCTGCTGGATGGAATCCTCTTCTATGTTGTCATTTCTCACTGCTTGTTCCTGA 3421  
Db 2706 TCGCATCCAGGA---AACAGAGCGGCTATCCGCGATCCATGCCCCCGAATGCAAG 2762  
Qy 3422 GAGTGAAGTTTCAAGAGGCGCAGAGCCCGCGGTACTACAGAGGGCCAGAACCAAGCTCT 3481  
Db 2763 AGTGGGAGGCAACGATGGCGCTTTGTTGTCGACCGGACGGGACGCTCTCGCTGATAC 2822  
Qy 3482 ATAAAGAGCTCAATCTTAGGACGAAGAGAGGATGAGTGTGTTTGGACMAAGAGAGTGGCC 3541  
Db 2823 AGAAAGAAATTTGCTTTCAGGCACTCTCATGAGTGTCTT-----CCCGTTTTTC 2869  
Qy 3542 GGGACCTTGAGTGGGGGGAAGCCGAGAGGAAGAACCTCTCAGGAAGGCTGTACAAATG 3601  
Db 2870 CGCTGAGGTCACTGCGTGTGATGAGCGCTGGCGCTGCTCGCGACGCGCAG-----CTG 2925  
Qy 3602 AACTGCAAGAAAGTAAGATGGCGGAGCCTACAGTGAGATTGGGATGAAAGCGCAGCGCC 3661  
Db 2926 CTCACCACTCGAGGGCGTGCAGCGCTGCAGAGCCGAGTCAGAACTGC----- 2978  
Qy 3662 GGAGGGGGAAGGGCAAGTGGCCCTTTACAGGTCTCTAGTACAGCCCAAGAGACACT 3721  
Db 2979 -----TCCAAAGGACCTCAAGGCTTCCGAGGACACTAGGCTCACTCCATCGAGCCAG 3033  
Qy 3722 ACAGCGCCTTCACTGACGGCCCTGCCCCCTCGCTTAACCTGACGCGCCGCGGATCCGG 3781  
Db 3034 TGTA-----GAGATAAGCTTTATCG 3052  
Qy 3782 ATTAGTCCAAATTTGTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTGACTCAACA 3841  
Db 3053 ATTAGTCCAAATTTGTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTGACTCAACA 3112  
Qy 3842 ATATCAAGCTGAAGCTTATAGAGTACGAGCCATAGATAAATAAAGATTTTATTAG 3901  
Db 3113 ATATCAAGCTGAAGCTTATAGAGTACGAGCCATAGATAAATAAAGATTTTATTAG 3172  
Qy 3902 TCTCCAGAAAAGGGGGAATGAAAGACCCACCTGTAGTTTGGCAAGCTAGCTTAAGT 3961  
Db 3173 TCTCCAGAAAAGGGGGAATGAAAGACCCACCTGTAGTTTGGCAAGCTAGCTTAAGT 3232  
Qy 3962 AACCCCAATTTTGAAGCATGG-AAAATACATAACTAGATAAGAGATTTAGATCAAG 4020  
Db 3233 AACCCCAATTTTGAAGCATGGAATAATACATACTGAGATAGAGAGTTTCAATCAAG 3292  
Qy 4021 GTTAGGAACAGA-GAGACAGCAGAAATATGGGCCAAAACAGGATATCTGTGGTAAGCAGTTC 4079  
Db 3293 GTCAGGAACAGATGGAACAGCTGAATATGGGCCAAAACAGGATATCTGTGGTAAGCAGTTC 3352  
Qy 4080 CTGCCCC-GCTCAGGGCCAGAAACAGTTGGAACAGGAGATATGGGCCAAAACAGGATATC 4138  
Db 3353 CTGCCCCGGCTCAGGGCCAGAAACAGATGGAACAGCTGAATATGGGCCAAAACAGGATATC 3412



```
Db 5181 AGGGCTTACCATCTGGCCCGAGTGTGCAATGATACCGCGAGACCCACGCTCACCGGCTC 5240
QY 6358 CAGATTATACGAAATAAACAGCCAGCCGGAAGCCGCGCAGAGAGTGGTCTCTGCAA 6417
Db 5241 CAGATTATACGAAATAAACAGCCAGCCGGAAGCCGCGCAGAGAGTGGTCTCTGCAA 5300
QY 6418 CTTTATCCGCTCCATCCAGTCTATTAAATTGTTGCGGGAAGCTAGAGTAAGTAGTTCCG 6477
Db 5301 CTTTATCCGCTCCATCCAGTCTATTAAATTGTTGCGGGAAGCTAGAGTAAGTAGTTCCG 5360
QY 6478 CAGTTAATAGTTTGGCCAAAGTTGTTGCAATTTGCTTACAGGC-TGCTGGTGTCAAGCTCGT 6536
Db 5361 CAGTTAATAGTTTGGCCAAAGTTGTTGCAATTTGCTTACAGGCATCTGGTGTCAAGCTCGT 5420
QY 6537 CGTTTGGTATGGCTTTCATTCAGCTCCGGTTCCTCAACGATCAAGCGAGTTACATGATCCC 6596
Db 5421 CGTTTGGTATGGCTTTCATTCAGCTCCGGTTCCTCAACGATCAAGCGAGTTACATGATCCC 5480
QY 6597 CCATGTTGTGCAAAAAGCGTTAGCTCCTTCGGTCTCCGATCGTGTGTGCAAGTAAGT 6656
Db 5481 CCATGTTGTGCAAAAAGCGTTAGCTCCTTCGGTCTCCGATCGTGTGTGCAAGTAAGT 5540
QY 6657 TGGCCGCAAGTGTATCACTCATGTGTTATGGCAGCACTGCATAAATCTCTTACTCTCATGC 6716
Db 5541 TGGCCGCAAGTGTATCACTCATGTGTTATGGCAGCACTGCATAAATCTCTTACTCTCATGC 5600
QY 6717 CATCCGTAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCAAGTCAATCTTGAGAATAGT 6776
Db 5601 CATCCGTAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCAAGTCAATCTTGAGAATAGT 5660
QY 6777 GTATGCGCGGACCGAGTTGCTCTTGGCCGCGGTCAATAACGGGATAAATACCGGCCACATA 6836
Db 5661 GTATGCGCGGACCGAGTTGCTCTTGGCCGCGGTCAACACGGGATAAATACCGGCCACATA 5720
QY 6837 GCAGAACTTTAAAGTGTCTCATCTGAAACGTTCTTCGGGGCGAAGAACTCTCAAGGA 6896
Db 5721 GCAGAACTTTAAAGTGTCTCATCTGAAACGTTCTTCGGGGCGAAGAACTCTCAAGGA 5780
QY 6897 TCTTACCGCTGTTGAGATCCAGTTTCGATGTAAACCCACTCTGTCACCCCACTGTCTTCAG 6956
Db 5781 TCTTACCGCTGTTGAGATCCAGTTTCGATGTAAACCCACTCTGTCACCCCACTGTCTTCAG 5840
QY 6957 CATCTTTTACTTTCACGAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAA 7016
Db 5841 CATCTTTTACTTTCACGAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAA 5900
QY 7017 AAAAGGGAATAAGGCGACACGGAATGTTGAATACTACTACTCTCTCTCTTTTCAATATT 7076
Db 5901 AAAAGGGAATAAGGCGACACGGAATGTTGAATACTACTACTCTCTCTCTTTTCAATATT 5960
QY 7077 ATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTTATTAGA 7136
Db 5961 ATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTTATTAGA 6020
QY 7137 AAAATAAACAATAAGGGTTCGCGCACATTTCCCGAAAAGTCCCACTGACGTCTTAAG 7196
Db 6021 AAAATAAACAATAAGGGTTCGCGCACATTTCCCGAAAAGTCCCACTGACGTCTTAAG 6080
QY 7197 AAACCATTTATCATGACATTAACCTATAAAATAGGCGTATCACGAGGCCCTTTTCGTC 7256
Db 6081 AAACCATTTATCATGACATTAACCTATAAAATAGGCGTATCACGAGGCCCTTTTCGTC 7256
QY 7257 T 7257
Db 6141 T 6141
```



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2005, 19:41:59 ; Search time 289 Seconds  
(without alignments)  
592.854 Million cell updates/sec

Title: US-10-006-771b-2

Perfect score: 2350

Sequence: 1 MGWSCIILFLVATATGVHSD.....LSTATKDTYDALHMQALPPR 443

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq\_16Dec04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2350	100.0	443	6	ABG76488 Humanised
2	2350	100.0	443	6	ABG74240 Chimaeric
3	1590.5	67.7	473	2	AAW26646 Chimeric
4	1590.5	67.7	514	2	AAW26647 Chimeric
5	1447	61.6	532	2	AAW73051 Z33dCR2.z
6	1409.5	60.0	643	2	AAW73050 Z33g2G237
7	1406.5	59.9	651	2	AAW26649 Chimeric
8	1406.5	59.9	692	2	AAW26650 Chimeric
9	1304.5	55.5	643	2	AAW73048 A33 chime
10	1228	52.3	634	6	ABB82300 CD19:zeta
11	1198	51.0	633	3	AAW84965 Amino aci
12	1094	46.6	444	2	AAW36845 Single ch
13	1038	44.2	443	2	AAW24025 Single ch
14	1015.5	43.2	461	2	AAW85508 Leader-sc
15	991	42.2	523	3	AAW44994 HD70scFv-
16	986	42.0	524	3	AAW44995 HD70scFv-
17	966	41.1	428	2	AAW24027 Single ch
18	965	41.1	352	2	AAW06272 Anti-Fc g
19	937.5	39.9	574	7	AAW52216 E3B1 prot
20	937.5	39.9	2606	7	AAW52217 pGLEN-EH3
21	937.5	39.9	2606	7	AAW52214 pGLEN-EH3
22	920	39.1	411	8	ADR46829 Human pB1
23	918	39.1	483	2	AAW82315 Chimeric
24	910	38.7	631	7	ADH34696 CE7-speci
25	908	38.6	403	2	AAW26648 Chimeric

ALIGNMENTS

RESULT 1

ABG76488

ID ABG76488 standard; protein; 443 AA.

AC ABG76488;

XX

XX

DT 13-MAY-2003 (first entry)

XX

DE Humanised anti-CEA antibody, hMN14.

XX

XX

KW CEA; carcinoembryonic antigen; IgTCR; T-cell receptor; cancer; tumour;

KW colorectal cancer; breast cancer; lung cancer; hMN14; cytostatic; mouse;

KW human; zeta signalling chain; CD8alpha hinge; humanised antibody.

XX

OS Homo sapiens.

OS Mus sp.

OS Synthetic.

OS Chimeric.

XX

PN US2002165360-A1.

PD 07-NOV-2002.

XX

PF 10-DEC-2001; 2001US-00006771.

XX

PR 30-NOV-2000; 2000US-0250087P.

PR 30-NOV-2000; 2000US-0250090P.

XX

XX

PA (JUNG/) JUNGHANS R P.

XX

PI Junghans RP;

XX

XX

DR WPI; 2003-298705/29.

DR N-PSDB; ABX13168.

XX

PS Disclosure; Page 7-8; 20pp; English.

XX

XX

CC The invention relates to a chimaeric molecule comprising the

CC carcinoembryonic antigen (CEA) binding domain of humanised antibody MN14

CC as a single chain antibody with a (GGSGS)<sub>3</sub> linker, the zeta signaling

CC chain of the T cell receptor (TCR) and an intervening CD8alpha hinge in

CC which the cysteine residues have been mutated, with the IgTCR molecule

CC occupying nucleotides 2426-3766 of the retroviral vector sequence.

CC appearing as ABX13168. The new chimaeric molecule expressed in T cells,

CC

CC



DB 421 YQGLSTATKDYDALHMQALPPR 443

|||||

AAW26646

RESULT 3

ID AAW26646 standard; protein; 473 AA.

XX AAW26646;

XX

DT 17-OCT-2003 (revised)

DT 12-FEB-1998 (first entry)

XX

DE Chimeric receptor hCTMO1/CD8/zeta.

XX

KW Cell activation; chimeric receptor; DNA delivery; CTMO1; scFv; CD8;

KW T cell receptor zeta chain; cell proliferation; cytokine; inflammation;

KW effector; cell differentiation; antibody secretion; phagocytosis;

KW tumour infiltration; adhesion; infection; cancer; allergy;

KW rheumatoid arthritis; osteoarthritis; asthma; eczema;

KW inflammatory bowel disease; cystic fibrosis; sickle cell anaemia;

KW psoriasis; multiple sclerosis; transplant rejection; diabetes;

KW graft versus host disease; human; therapy.

XX

OS Homo sapiens.

OS Chimeric.

XX

PN WO9723613-A2.

XX

PD 03-JUL-1997.

XX

XX 23-DEC-1996; 96WO-GB003209.

XX

XX 21-DEC-1995; 95GB-00026131.

XX

PA (CLLT ) CELLTECH THERAPEUTICS LTD.

XX

XX Bebbington CR, Lawson ADG, Weir ANC, Finney HM;

XX

XX WPI; 1997-351052/32.

DR N-PSDB; AAT90509.

DR

XX

PT New DNA systems for activating cells - comprising DNA coding for a

PT chimeric receptor comprising 2 or more different cytoplasmic signalling

PT components.

XX

PS Disclosure; Fig 4; 90pp; English.

XX

CC This protein comprises a chimeric receptor consisting of an scFv

CC engineered from anti-CD3 human antibody CTMO1 linked to an extracellular

CC spacer in the form of part of human CD8 hinge, linked to the

CC extracellular, transmembrane and intracellular components of the human T

CC cell receptor zeta chain. It can be expressed in host cells (e.g. Jurkat)

CC using a chimeric receptor gene (see AAT90509) constructed from DNA

CC cassettes of each component. In a claimed cell activation process an

CC effector cell is transformed with DNA encoding a chimeric receptor

CC containing 2 or more different cytoplasmic signalling components. Also

CC claimed is use of DNA encoding a recombinant chimeric receptor in a DNA

CC delivery system. The DNA delivery systems can be used for the activation

CC of cells to provide e.g. an increase in cell proliferation, expression of

CC cytokines with e.g. pro- or anti-inflammatory responses, stimulation of

CC cytolytic activity, differentiation or other effector functions, antibody

CC secretion, phagocytosis, tumour infiltration and/or increased adhesion.

CC They can be used in the treatment of e.g. infectious disease,

CC inflammatory disease, cancer, allergic/atopic disease, congenital

CC disease, dermatologic disease, neurologic disease, transplants and

CC metabolic/idiopathic disease (claimed). In particular, they can be used

CC in the treatment of rheumatoid arthritis, osteoarthritis, inflammatory

CC bowel disease, asthma, eczema, cystic fibrosis, sickle cell anaemia,

CC psoriasis, multiple sclerosis, organ or tissue transplant rejection,

CC graft-versus-host disease or diabetes (claimed). (Updated on 17-OCT-2003

CC to standardise OS field)

XX

SQ Sequence 473 AA;

Query Match 67.7%; Score 1590.5; DB 2; Length 473;

Best Local Similarity 66.0%; Pred. No. 6.2e-102;

Matches 318; Conservative 42; Mismatches 63; Indels 59; Gaps 8;

QY 6 IILFLVATATGVHSDIQLTQSPSLASVGDRTVITCKASQDV-----GTSVANYQKPG 60

DB 7 VLGLLLWLTDARCDIQMTQSPSTLSASVGDRTVITCRSSKSLHSHGDTFLYWFQKPG 66

QY 61 KAPKLLIYWTSTRHTGVPSRFGSGSGCTDFTFTISSIQLPEDIATYVCOQVSLYR-SFQOG 119

DB 67 KAPKLLMYRMSNLASVPSRFGSGSGTEFTLTITSSLPDDFATYICMQLHEYFTFCQG 126

QY 120 TKVEIKR-----GGSGSGSGSGSGSEVQLVESGGGVVQVQPSRLSLSCSAGF 168

DB 127 TKVEVKETGGSGSGSGSGSGSGSGSGGSGIQLVQSGAEVKFGSSVKVSKASGY 186

QY 169 DFTTYMWSVRQAPGKLEWIGEIHDPDSSTINVAPSLKDRFTISRDNAKNTFLQMDSLR 228

DB 187 TFDYIINMRQAPGQGLEWIGWIDPGSGNTKYNEKFKGRATLTVDSTNTAYMELSSLR 246

QY 229 PEDTGVVFCFCA---SLYFGFPWFAYWGQGTPTVTSS-----A 261

DB 247 SEDTAFYFCAREKTTYIYAMD---YWGQGLTVTVSSASTKGPTNSIMYFSHFVFLPA 303

QY 262 KPTTTPAPRPTPTAPTIASQPLSLRPEAARPAAGGAVHTRGLDFALDPLKCLYLDGILFI 321

DB 304 KPTTTPAPRPTPTAPTIASQPLSLRPEA-----QSFGL---LQPKLYLDGILFI 351

QY 322 YGVILTALFLRVKFSRABPPAYQOQGNQLYNELNGRREYDVLDRKRGRDPEWGKPR 381

DB 352 YGVILTALFLRVKFSRABPPAYQOQGNQLYNELNGRREYDVLDRKRGRDPEWGKPR 411

QY 382 RKNPQEGLYNELOKQWAEAYSEIGMKGERRRGHGLYQGLSTATKDYDALHMQALP 441

DB 412 RKNPQEGLYNELOKQWAEAYSEIGMKGERRRGHGLYQGLSTATKDYDALHMQALP 471

QY 442 PR 443

DB 472 PR 473

RESULT 4

AAW26647

ID AAW26647 standard; protein; 514 AA.

XX AAW26647;

AC

XX

DT 17-OCT-2003 (revised)

DT 12-FEB-1998 (first entry)

XX

DE Chimeric receptor hCTMO1/CD8/zeta-CD28.

XX

KW Cell activation; chimeric receptor; DNA delivery; CTMO1; scFv; CD8; CD28;

KW T cell receptor zeta chain; cell proliferation; cytokine; inflammation;

KW effector; cell differentiation; antibody secretion; phagocytosis;

KW tumour infiltration; adhesion; infection; cancer; allergy;

KW rheumatoid arthritis; osteoarthritis; asthma; eczema;

KW inflammatory bowel disease; cystic fibrosis; sickle cell anaemia;

KW psoriasis; multiple sclerosis; transplant rejection; diabetes;

KW graft versus host disease; human; therapy.

XX

OS Homo sapiens.

OS Chimeric.

XX

PN WO9723613-A2.

XX

PD 03-JUL-1997.

XX

XX 23-DEC-1996; 96WO-GB003209.

XX

XX 21-DEC-1995; 95GB-00026131.

XX

XX	(CLLT ) CELLTECH THERAPEUTICS LTD.	382	RKNPQEGLYNELQDKMAEAYSEIGMKGERRGKHDGLYQGLSTATKDTYDALHMQALP	441
XX	Bebington CR, Lawson ADG, Weir ANC, Finney HM;	412	RKNPQEGLYNELQDKMAEAYSEIGMKGERRGKHDGLYQGLSTATKDTYDALHMQALP	471
XX	WPI; 1997-351052/32.			
XX	N-PSDB; AAT90510.			
PT	New DNA systems for activating cells - comprising DNA coding for a			
PT	chimeric receptor comprising 2 or more different cytoplasmic signalling			
PT	components.			
PS	Disclosure; Fig 5; 90pp; English.			
XX	This protein comprises a chimeric receptor consisting of an scFv			
CC	engineered from anti-CD3 human antibody CTMO1 linked to an extracellular			
CC	spacer in the form of part of human CD8 hinge, linked to the			
CC	extracellular, transmembrane and intracellular components of the human T			
CC	cell receptor zeta chain, fused to the intracellular component of human			
CC	CD28. It can be expressed in host cells (e.g. Jurkat) using a chimeric			
CC	receptor gene (see AAT90510) constructed from DNA cassettes of each			
CC	component of the receptor. In a claimed cell activation process an			
CC	effector cell is transformed with DNA encoding a chimeric receptor			
CC	containing 2 or more different cytoplasmic signalling components. Also			
CC	claimed is use of DNA encoding a recombinant chimeric receptor in a DNA			
CC	delivery system. The DNA delivery systems can be used for the activation			
CC	of cells to provide e.g. an increase in cell proliferation, expression of			
CC	cytokines with e.g. pro- or anti-inflammatory responses, stimulation of			
CC	cytolytic activity, differentiation or other effector functions, antibody			
CC	secretion, phagocytosis, tumour infiltration and/or increased adhesion.			
CC	They can be used in the treatment of e.g. infectious disease,			
CC	inflammatory disease, cancer, allergic/atopic disease, congenital			
CC	disease, dermatologic disease, neurologic disease, transplants and			
CC	metabolic/idiopathic disease (claimed). In particular, they can be used			
CC	in the treatment of rheumatoid arthritis, osteoarthritis, inflammatory			
CC	bowel disease, asthma, eczema, cystic fibrosis, sickle cell anaemia,			
CC	psoriasis, multiple sclerosis, organ or tissue transplant rejection,			
CC	graft-versus-host disease or diabetes (claimed). (Updated on 17-OCT-2003			
CC	to standardise OS field)			
XX	Sequence 514 AA;			
XX	Query Match 67.7%; Score 1590.5; DB 2; Length 514;			
XX	Best Local Similarity 66.0%; Pred No. 6.9e-102;			
XX	Matches 318; Conservative 42; Mismatches 63; Indels 59; Gaps 8;			
QY	6 IILFVATATGTVHSDIQTQSPSSLSASVGRVITTCASQDV-----GTSVAMYQKQPG 60			
Db	7 VLGLLLWLTDARCDIQMTQSPSTLSASVGRVITTCRSSKLLSHNSGDTFLYWFQKPG 66			
QY	61 KAPKLLIYWTSTRHTGVPSRSGSGCTDFTTISLOPEDIATYTCQOYSLYR-SFQGG 119			
Db	67 KAPKLLMYRMNLSAGVPSRSGSGSGTEFTLTISLQDPDPATYTCMQLHLEYPFTQGG 126			
QY	120 TKVEIKR-----GGSGGGSGGGSGGVQVPGVGRSLRLSCSASGF 168			
Db	127 TKVEIKRTGG 186			
QY	169 DFTTYWNSVRQAPKGLEWIGEIHDPDSTTINYAPSLKDRFTISRDNKATFLFQMSLR 228			
Db	187 TFTDYINWMPAQPGLEWIGWDIPGSGNTKYNKFKGRATLTVDSTNTAYMELSLR 246			
QY	229 PEDTGVIYCA-----SLYGFPPFAFWGGGTPTVSS-----A 261			
Db	247 SEDTAFYFCAREKTYTYAMD---YWGQGTLTVTSSASTKGPTNSIMYFSHFVFLPA 303			
QY	262 KPTTTPAPRPTTAPATIASQPLSLRPEAARPAAGCAVHTGLDPAIDPKLCVLLDGLFI 321			
Db	304 KPTTTPAPRPTTAPATIASQPLSLRPEA-----QSFGL---LDPKLCVLLDGLFI 351			
QY	322 YGVILTALFLRVKFSRSAPPAYQOQGNQLYNELMGRREEYVLDKRRGRDPENGGKPR 381			
Db	352 YGVILTALFLRVKFSRSADAPAYQOQGNQLYNELMGRREEYVLDKRRGRDPENGGKPR 411			
QY	382 RKNPQEGLYNELQDKMAEAYSEIGMKGERRGKHDGLYQGLSTATKDTYDALHMQALP 441			
Db	412 RKNPQEGLYNELQDKMAEAYSEIGMKGERRGKHDGLYQGLSTATKDTYDALHMQALP 471			
QY	442 PR 443			
Db	472 PR 473			
RESULT 5				
AAW73051				
ID	AAW73051 standard; protein; 532 AA.			
XX	AAW73051;			
AC	17-OCT-2003 (revised)			
DT	02-FEB-1999 (first entry)			
XX	Z33dCH2.z chimeric receptor.			
XX	Chimeric receptor; A33; Z33; colorectal cancer; antigen; tumour;			
KW	cytolysis; gene therapy; signal transduction; monoclonal antibody;			
KW	single chain antibody; scAb; humanised antibody; human; Z33dCH2.z.			
XX	Mus sp.			
OS	Chimeric.			
OS	Location/Qualifiers			
FT	Key			
FT	Peptide			
FT	Region			
FT	Peptide			
FT	Domain			
FT	Domain			
FT	Domain			
FT	Region			
FT	Region			
XX	WO9841613-A1.			
PD	24-SEP-1998.			
PF	13-MAR-1998; 98WO-US003797.			
XX	14-MAR-1997; 97US-00815030.			
XX	(OTTE/) OTTEN G R.			
PA	(GREE/) GREENBURG G B.			
PA	(CASE/) CASENTINI-BOROCZ D.			
PA	(FINE/) FINER M H.			
XX	Otten GR, Greenburg GB, Casentini-Borocz D, Finer MH;			
XX	WPI; 1998-521211/44.			
DR	N-PSDB; AAV58937.			
XX	New chimeric DNA encoding membrane-bound receptor for non-MHC restricted			
PT	signal activation - inducing cytotoxic effector cells or cytokine			
PT	production in presence of tumour cells expressing the A33 antigen, and			
PT	related vectors.			









```

QY 336 SRSAPPPAYQQQONQYLNELNIGRREYDVLDKRRGRDPDMGKPRKRNPOEGLYNELQK 395
Db 544 SRSADAPAYQQQONQYLNELNIGRREYDVLDKRRGRDPDMGKPRKRNPOEGLYNELQK 603
QY 396 DKQAEAYSEIGMKGRRCKGHDGLYQGLSTATKDTYDALHMQALPPR 443
Db 604 DKQAEAYSEIGMKGRRCKGHDGLYQGLSTATKDTYDALHMQALPPR 651

RESULT 9
AAW73048
ID AAW73048 standard; protein; 643 AA.
XX AC
XX AAW73048;
DT 17-OCT-2003 (revised)
DT 02-FEB-1999 (first entry)
XX DE
XX A33 chimeric receptor.
KW Chimeric receptor; A33; colorectal cancer; antigen; tumour; cytolysis;
KW gene therapy; signal transduction; monoclonal antibody; mouse.
XX OS
XX Mus sp.
XX Homo sapiens.
XX Chimeric.
FH Key
FT Peptide
FT 1..22
FT /label= Sig peptide
FT /note= "from V kappa chain of human Ab 96-6"
FT Region
FT 23..129
FT /label= A33-VL
FT Peptide
FT 130..147
FT /label= Linker
FT Region
FT 148..264
FT /label= A33_VH
FT Domain
FT 265..490
FT /label= Constant
FT /note= "human Ig2 constant domain (Cgamma2-1 domain
FT deleted, with G237A mutation to decrease affinity for
FT human IgG Fc receptors)"
FT Region
FT 491..507
FT /label= M1
FT /note= "human IgG2 M1 segment"
FT Domain
FT 508..531
FT /label= Transmembrane
FT /note= "human CD4 transmembrane domain"
FT Region
FT 532..643
FT /label= Zeta
FT /note= "intracellular portion of human TCR-associated
FT zeta chain"
XX WO9841613-A1.
XX 24-SEP-1998.
XX 13-MAR-1998; 98WO-US003797.
XX 14-MAR-1997; 97US-00815030.
XX (OTTE/) OTTEN G R.
XX (GREE/) GREENBURG G B.
XX (CASE/) CASENTINI-BOROCZ D.
XX (FINE/) FINER M H.
XX Otten GR, Greenburg GB, Casentini-Borocz D, Finer MH;
XX WPI; 1998-521211/44.
XX DR N-PSDB; AAV58929.
XX New chimeric DNA encoding membrane-bound receptor for non-MHC restricted
XX signal activation - inducing cytotoxic effector cells or cytokine

```

```

PT production in presence of tumour cells expressing the A33 antigen, and
PT related vectors.
XX Example 3; Page 41; 90pp; English.
XX This is the amino acid sequence of a murine A33 chimeric receptor that
XX comprises a signal peptide from the V kappa chain of human antibody 98-6,
XX light chain and heavy chain variable regions of anti-colon cancer antigen
XX A33 murine monoclonal antibody A33, human IgG2 constant domain and M1
XX segment, human CD4 transmembrane domain and an intracellular portion of
XX human TCR-associated zeta chain. A humanised version of the A33 chimeric
XX receptor has been constructed (see AAW73049). An expression construct
XX (see AAV58929) encoding the chimeric receptor is provided. Novel chimeric
XX proteins, and DNA encoding them, are claimed, in which the chimeric
XX proteins have an extracellular domain capable of binding to A33 in a non-
XX MHC restricted manner, a transmembrane domain, and a cytoplasmic domain
XX capable of activating a signalling pathway. Binding of A33 to the
XX extracellular domain results in transduction of a signal and activation
XX of a signalling pathway in the cell, such that the cell may be induced to
XX carry out various functions relating to the signalling pathway. The
XX chimeric DNA may be used to modify lymphocytes as well as haematopoietic
XX stem cells as precursors to a number of important cell types. The
XX modified cells can be transplanted (especially as a bone marrow
XX transplant) into a mammal to provide a source of cytotoxic effector cells
XX that can kill cells expressing the tumour-associated A33 antigen and
XX cytokine producing cells. The modified cells are activated in presence of
XX A33+ cells. Expression constructs and retrovirus vectors are provided.
XX (Updated on 17-OCT-2003 to standardise OS field)
XX Sequence 643 AA;
XX
Query Match 55.5%; Score 1304.5; DB 2; Length 643;
Best Local Similarity 46.7%; Pred. No. 5.9e-82;
Matches 293; Conservative 41; Mismatches 92; Indels 201; Gaps 13;
QY 16 GVHSDIQLTQSPSLASVGDRTVITCKASQDVGTSAVYQOKPKKLIYWTSTRT 75
Db 19 GARCDDVMTQSKFMSTVSGDRVSTICKASQNVRTVAWYQOKPQSKPTLIYLASNRT 78
QY 76 GVPFRSGSGSGTDTFTTISLQEDATATYTC-QQSYLSYRSFGQTKVEIKRGGSGSG-- 132
Db 79 GVPDRFTGSGSGTDTFTLTISNVQSEDLADYFCLOHWSYPLTFGSGTKLEVKGSTSGSKP 138
QY 133 GSGSGSGSEVOLVESGGVQPCRSRLRLSCASGDFFTTYWMSVROAPCKGLEWIGE1 192
Db 139 GSGEGSTGKEVGLVESGGGLVKPGSLKSCAASGFAPSTYDMSWVRVTPKRLWVAWI 198
QY 193 HPDSSTINYAPSLKDRFTISRDNKNTLFLQMDSLRPEDTGVYFCASLYFGFPWFAYWGQ 252
Db 199 SSGGSYTYLDSVKGRTISRDSARNTLYLQMSLSRSEDATLYYCPTTV-VP-FAYWGQ 256
QY 253 GTPVTVSSAKPTTT---PAPRPPTAPTIASQP-----LSLRPEARPA----- 294
Db 257 GTLVTVSSERKCCVECPPCPAPPAAPSVFLFPKPKDTLMISRTPEVTVVVDVSHEDP 316
QY 295 -----GCAGVH-----TRGL----- 303
Db 317 EVQFNWYVDGMEVHNATKPREQPNSTFRVSVLTVVHQDWLNGKEYKCKVSNKGLPAP 376
QY 304 ----- 313
Db 377 IEKTSKTKGQPREPQVYTLPPSREMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNY 436
QY 314 LL-----DGILFIY----- 322
Db 437 KTTTPMLDSDGSEFFLYSKLTVDKSRWQQGNVFCVMHEALHNYHTQKSLSPQLLEE 496
QY 323 -----GVILTALFL-----RVKFSASPPPAQQOONLYNELN 356
Db 497 SCASAQDGEDMALIVLGGVAGLLLF1GLG1FFCVRVKFSASADAPAYQQQONQYNELN 556
QY 357 LGREEYDVLDKRRGRDPDMGKPRKRNPOEGLYNELQKQWABAYSIGMKGRRRKGK 416

```





CC from these PCR products. Cells expressing recombinant TCR are used to  
 CC identify antigens associated with a tumour and to treat tumours in  
 CC humans. Transgenic mice are a more convenient source of CTL than the  
 CC tumour-infiltrating lymphocytes previously used. TCR can be humanised to  
 CC reduce side-reactions and short peptide derivatives of TCR are more  
 CC economical to produce than TCR itself, particularly when expressed as a  
 CC single-chain molecule rather than as a dimer  
 XX  
 SQ Sequence 444 AA;

Query Match 46.6%; Score 1094; DB 2; Length 444;  
 Best Local Similarity 54.7%; Pred. No. 1.5e-67;  
 Matches 248; Conservative 35; Mismatches 106; Indels 64; Gaps 12;

Qy 22 QLTQSPSSLSASVGRVTITCKASQDVGTIS-VAMVQQKPGKAPKLLIYWTSTRTHTGVPSSR 80  
 Db 25 QVQGSASLVLEGENAEALQSFSS--IFTNQVQVQRPGRGLVSLLY-----NPSG 74  
 Qy 81 FSGSGSGTDFT-----FTISLQPEDIAITYYQQVYS-----LYRSGQGTKEIKRGG 128  
 Db 75 TKQSGRLTSTTVIKERRSSLSHSSQITDSGYLCSNSGGSNKLTFGKTKLSVKSGG 134  
 Qy 129 SGSGSGSGSGSGSEVQLVES--GGGVQVQPSRLRLSCSASGDFDTTYWMSWVRQAPGKGLE 187  
 Db 135 GSGSGSGSGSGGSEAAVTQSPRNKAVTGGKVTLSQNTNNHNMV---WYRQDTGHGLR 191  
 Qy 188 WIGIEHPDSTINYPASLKDRTISRDNKYN-TLFLQMDSLRPEDTGYVFCASLYFGFPW 246  
 Db 192 LIHYSYAGST--BKGDIPDGYKASRPSQENFSLILELAT--PSQTSYVFCASGETGNE 247  
 Qy 247 FAYWQGTPTVTVSS-----AKPTTTPAPRPTPAPTPTIAPISQPLSLRPEAA 290  
 Db 248 RLFFGHGHTKLSVLTNSIMYFHFVFPVFLPAKPTTTPAPRPTTPTIAPISQPLSLRPS 307  
 Qy 291 RPAAGGAVHTRGDLDPKLCYLLDGLFLTYGVILTALFLRVKFSRSPAPPAVQQGQ 350  
 Db 308 R-----DPLCYLLDGLFLTYGVILTALFLRVKFSRSPAPPAVQQGQ 351  
 Qy 351 LYNELNLGRREYDVLDRGRDPEMGKPKRRKNPQGLYNELQDKWAEYSIGNKGE 410  
 Db 352 LYNELNLGRREYDVLDRGRDPEMGKPKRRKNPQGLYNELQDKWAEYSIGNKGE 411  
 Qy 411 RRRKGHDGLYQGLSTATKOTYDYLHMQALPPR 443  
 Db 412 RRRKGHDGLYQGLSTATKOTYDYLHMQALPPR 444

RESULT 13  
 AAW24025  
 ID AAW24025 standard; protein; 443 AA.  
 XX  
 AC AAW24025;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 04-MAR-1998 (first entry)  
 XX  
 XX Single chain antigen hybrid receptor.  
 DE  
 XX  
 KW Hybrid receptor; single chain antigen; gene therapy; diagnosis;  
 KW signal conduction; receptor; control region.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 PH Location/Qualifiers  
 FT 1..19  
 FT /label= leader\_peptide  
 FT 20..444  
 FT /note= "single chain antigen hybrid receptor"  
 FT 20..139  
 FT /label= VH B\_1-8  
 FT /note= "variable heavy chain region of B1-8 antibody"  
 FT 140..154  
 FT Region

FT /label= (Gly\_Ser)3  
 FT /note= "linker region"  
 FT 155..264  
 FT /label= VL B\_1-8  
 FT /note= "variable light chain region of B1-8 antibody"  
 FT 265..309  
 FT /label= CD\_8-alpha\_hinge  
 FT 310..330  
 FT /label= transmembrane\_domain  
 FT 331..444  
 FT /label= cytoplasmic\_domain  
 XX  
 XX W09720938-A2.  
 XX  
 XX 12-JUN-1997.  
 XX  
 XX 03-DEC-1996; 96WO-DE002334.  
 XX  
 XX 05-DEC-1995; 95DE-01045351.  
 XX  
 XX (UYFR-) UNIV FREIBURG KLINIKUM ALBERT-LUDWIGS.  
 XX  
 XX Mertelsmann R, Kulmburg P, Rosenthal F;  
 XX  
 XX WPI; 1997-319784/29.  
 XX  
 XX N-PSDB; AAT77137.  
 XX  
 XX Cells with hybrid receptor having extracellular and intracellular regions  
 XX of different origins - useful in gene therapy and diagnosis of tumours.  
 XX  
 XX Example 3; Fig 4; 46pp; German.  
 XX  
 XX This sequence represents a novel single chain antigen hybrid receptor  
 XX (HR) and contains an extracellular domain specific for the hapten 4-  
 XX hydroxy-5-iodo-3-nitrophenyl acetate (NIP) coupled to the CD8-alpha  
 XX region and the transmembrane and signal-conducting intracellular parts of  
 XX the CD3-zeta molecule. Such hybrid receptors comprise a receptor part  
 XX localised on the outside of the cell and specific to a particular signal  
 XX molecule and a receptor part originating from another receptor, localised  
 XX on the inside of the cell and capable of setting off a signal inside the  
 XX cell. The cell should also contain at least one other gene construct with  
 XX a control region which can interact with the signal sent out by the  
 XX hybrid receptor and thereby control expression of a transgene bound to  
 XX this control region. Such cells are useful in gene therapy or for  
 XX diagnostic purposes. (Updated on 25-MAR-2003 to correct PR field.)  
 XX  
 SQ Sequence 443 AA;

Query Match 44.2%; Score 1038; DB 2; Length 443;  
 Best Local Similarity 51.2%; Pred. No. 1.2e-63;  
 Matches 243; Conservative 45; Mismatches 123; Indels 64; Gaps 19;

Qy 1 MGSCLILFLVATATGVSIDIQTQSPSSLSASVGRVTITCKASQDVGTIS--VAMVQQK 58  
 Db 1 MGSCLIMFLAATATGVSQVQLQQSGAEL-VKPGASVKLSCKASGYTFTSYMMHWKQR 59  
 Qy 59 PGKAPKLLIYWTSTRTHTGVPSSRSGSGTDF-----TFTI-----SSIQP 99  
 Db 60 PGRG----LEWI-----GRIDPNSGGTKYNEKFKSKATLVDFKSSYTMQLSLS 107  
 Qy 100 EDIATYCCQYSLYRS-----FGQGTKEIKRGGSGSGSGSGSEVQLVESGGGVVQ 154  
 Db 108 EDSAVVYCARVDYVGSYFDYWGQTTVTYSSGGSGSGSGSGGSGGSAVGTQSSALTS 167  
 Qy 155 PGRSLRLSC--SASGDFDTTYWMSWVRQAPG---KGLEWIGEIHDPDSSTINYPASLKDRT 210  
 Db 168 PGETVTLTKRSSTGAVTTSNYANWVQEKPDHLFTGL--IG-----GTNRAPGVPARFS 219  
 Qy 211 ISRDNAKNTLFLQMDSLRPEDTGYVFCASLYFGFPWYVQGTPTVTVSSAKPTTTPAPR 270  
 Db 220 GSLIGDKAA--LTITGAQTEDEAIYFCA--LWYNHNM--VFGGGTKLVLBEEFTTKVLR 274  
 Qy 271 PPTPA-PTIASQPLSLRPEAAPAAGGAVHTRGLDFALDPKLCYLLDGLFIYGVILTAL 329



Db	275	TPSPVHPTSP--ORPEDCRP--GSVKTGLDF--DPKLCYLLDGLFIYGVITLAL	328
Qy	330	FLRVFSRSAPPAYQOQGNQLYNELNLGRREEYDVLDRGRDPENGGK--PRKKNPQEG	388
Db	329	YLRRKFSRSAAETAAQLDPNQLYNELNLGRREEYDVLKRRARDPENGHGQORRRNPQEG	388
Qy	389	LYNELKOKOMAAEYSEIGKGGRRRGKHGDLGYGLGLSTATKTYDALHMQALPPR	443
Db	389	IYNALQKONMAAEYSEITGGERRRGKHGDLGYGLGLSTATKTYDALHMQTLAPR	443

RESULT 14  
AAR85508  
ID AAR85508 standard; protein; 461 AA.  
XX  
AC AAR85508;  
XX  
DT 16-MAR-1996 (first entry)

DE Leader-scFv(FRP5):lyt-2 hinge:zeta.

Single chain antibody; scFv; antibody engineering; antitumour;  
tumour antigen binding; T-cell receptor; cytotoxic T-lymphocyte;  
monoclonal antibody; erbB-2; cancer; cell targeting;  
adoptive immunotherapy.

OS Synthetic.

Key	Location/Qualifiers
PH	40..93
FT	/label= IGH_chain_leader
FT	94..819
FT	Domain
FT	/label= scFv(FRP) 5
FT	261..322
FT	Peptide
FT	/label= Lyt-2_hinge
FT	323..461
FT	Domain
FT	/label= Zeta_chain

PN WO9530014-A1.

09-NOV-1995.

20-APR-1995: 95WO-EP001494.

AA  
PR 02-MAY-1994: 94EP-00810244.

PA (CIBA ) CIBA GEIGY AG.

XX PI Groner B. Moritz D:

XX DB WPT: 1995-393085/50

DR N-PSDB; AAT05783.  
XX

PT New bifunctional proteins for use in killing tumour cells - contg. a  
PT tumour antigen binding domain, a hinge region and a zeta chain derived  
PT from a T-cell antigen receptor.

XX  
PS  
Example 1: Page 29-31: 46pp: English.

A bifunctional protein (AAR55505) consists of a single chain antibody, scFv(FRP15), directed against the tumour ErbB-2 antigen, a hinge region, CC and a functional zeta chain obd. from a T-cell receptor. The protein is CC expressed in host cells, esp. cytotoxic T-lymphocytes, providing them CC with a defined tumour cell specificity enabling targeting to defined CC tumour cells and MHC-unrestricted and MHC-independent tumour destruction CC in vitro or in vivo CC

Sequence 461 AA;

Query Match	43.2%	Score	1015.5	DB 2	Length	461
Best Local Similarity	52.0%	Pred. No.	4.4e-62			
Matches	225	Conservative	25	Mismatches	58	Gaps
				Indels	125	



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 26, 2005, 00:49:19 ; Search time 68 Seconds  
(without alignments)  
626.824 Million cell updates/sec

Title: US-10-006-771b-2  
Perfect score: 2350  
Sequence: 1 MGSCLILFLVATATGVHSD.....LSTATKDYDHALHMQALPPR 443

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: Piri:.\*  
2: Pir2:.\*  
3: Pir3:.\*  
4: Pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	723	30.8	163	2 A31768	T-cell receptor ze
2	609	25.9	164	2 A40104	T-cell receptor CD
3	605	25.7	166	2 JC4664	T-cell receptor ze
4	605	25.7	166	2 I46424	T-cell surface gly
5	521.5	22.2	188	2 A45089	T-cell receptor CD
6	511	21.7	177	2 S54817	T-cell receptor io
7	511	21.7	206	2 A35900	T-cell receptor CD
8	500.5	21.3	287	4 PC4402	peLB leader/Ig hea
9	488.5	20.8	119	1 AVMSX4	Ig heavy chain v r
10	486.5	20.7	119	1 AVMSJ5	Ig heavy chain v r
11	482	20.5	139	2 A25912	Ig heavy chain pre
12	481	20.5	130	2 S06817	Ig heavy chain v r
13	480.5	20.4	119	1 AVMS76	Ig heavy chain v r
14	475	20.2	118	2 A47329	Ig heavy chain v r
15	473.5	20.1	136	2 S13791	Ig heavy chain v r
16	465	19.8	117	1 G2MSU1	Ig heavy chain v r
17	457	19.4	121	2 A30560	Ig heavy chain v r
18	445.5	19.0	126	1 G1HUKL	Ig heavy chain v-I
19	445	18.9	117	1 G2MS73	Ig heavy chain v-I
20	443.5	18.9	108	1 K1HULY	Ig kappa chain v-I
21	441.5	18.8	123	2 S40331	Ig kappa chain - h
22	441	18.8	107	2 S36272	Ig lambda chain v
23	439.5	18.7	117	2 S46371	Ig kappa chain v r
24	438.5	18.7	129	2 S52789	Ig kappa chain v r
25	438	18.6	147	2 I37780	Ig variable region
26	437.5	18.6	131	2 S40352	Ig kappa chain v-J
27	436.5	18.6	117	2 S42466	Ig kappa chain v r
28	436.5	18.6	125	2 S40349	Ig kappa chain v-J
29	435.5	18.5	143	2 S23624	Ig heavy chain v r

30	435	18.5	124	2 S40336	Ig kappa chain v-J
31	434.5	18.5	108	2 I39154	Ig kappa chain (BR
32	434.5	18.5	122	2 S31117	Ig heavy chain - h
33	434.5	18.5	125	2 S40333	Ig kappa chain v-J
34	432	18.4	125	2 S30531	Ig heavy chain v r
35	431	18.3	116	1 HVMS44	Ig heavy chain pre
36	431	18.3	121	2 S19666	Ig heavy chain v r
37	430.5	18.3	110	2 S44118	Ig kappa chain v-J
38	430.5	18.3	127	2 S40367	Ig kappa chain v-I
39	429.5	18.3	108	1 K1HUAU	Ig kappa chain v-I
40	429.5	18.3	128	2 S46372	Ig light chain var
41	429	18.3	139	2 I37781	Ig variable region
42	429	18.3	140	2 S06816	Ig heavy chain pre
43	428.5	18.2	114	2 S46390	Ig heavy chain v r
44	428.5	18.2	117	2 S46376	Ig kappa chain v-J
45	428.5	18.2	122	2 S40370	Ig kappa chain - h

ALIGNMENTS

RESULT 1

A31768  
T-cell receptor zeta chain precursor - human  
C;Species: Homo sapiens (man)  
C;Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 09-Jul-2004  
C;Accession: A31768  
R;Weissman, A.M.; Hsu, D.; Orloff, D.G.; Modi, W.S.; Seunanez, H.; O'Brien, S.J.; Klausne  
Proc. Natl. Acad. Sci. U.S.A. 85, 9709-9713, 1988  
A;Title: Molecular cloning and chromosomal localization of the human T-cell receptor zet  
A;Reference number: A31768; MUID:89071765; PMID:2974162  
A;Accession: A31768  
A;Molecule type: mRNA  
A;Residues: 1-163 <WEI>  
A;Cross-references: UNIPROT:P20963; GB:J04132; NID:G623041; PIDN:AAA60394.1; PID:G623042  
C;Keywords: phosphoprotein; T-cell receptor; transmembrane protein  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;22-163/Product: T-cell receptor zeta chain #status predicted <MAT>

Query Match 30.8%; Score 723; DB 2; Length 163;  
Best Local Similarity 100.0%; Pred. No. 2.4e-40;  
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	307	LDPKLCYLDGILFIYGVILTAFLRVKFSRSAPPAQQQNLNGLNLRREYDVL	366
Db	27	LDPKLCYLDGILFIYGVILTAFLRVKFSRSAPPAQQQNLNGLNLRREYDVL	86
Qy	367	DKRGDDPEMGKPRRKNPQEGLYNELQDKMAEYSEIGMKGRRRGKHGDLGYQLST	426
Db	87	DKRGDDPEMGKPRRKNPQEGLYNELQDKMAEYSEIGMKGRRRGKHGDLGYQLST	146
Qy	427	ATKDTYDALHMQALPPR	443
Db	147	ATKDTYDALHMQALPPR	163

RESULT 2

A40104  
T-cell receptor CD3 zeta chain - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 20-Mar-1992 #sequence\_revision 20-Mar-1992 #text\_change 09-Jul-2004  
C;Accession: A40104; I55293  
R;Weissman, A.M.; Baniyash, M.; Hou, D.; Samelson, L.E.; Burgess, W.H.; Klausner, R.D.  
Science 239, 1018-1021, 1988  
A;Title: Molecular cloning of the zeta chain of the T cell antigen receptor.  
A;Reference number: A40104; MUID:88145643; PMID:3278377  
A;Accession: A40104  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-164 <WEI>  
A;Cross-references: UNIPROT:P24161; GB:M19729; NID:G201131; PIDN:AAA40171.1; PID:G201132  
R;Baniyash, M.; Hsu, V.W.; Seidin, M.F.; Klausner, R.D.  
J. Biol. Chem. 264, 13252-13257, 1989

A;Title: The isolation and characterization of the murine T cell antigen receptor zeta c  
A;Reference number: I55293; MUID:89327299; PMID:2787796  
A;Accession: I55293  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-164 <RES>  
A;Cross-references: GB:J04967; NID:g556326; PIDN:AAA50301.1; PID:g556327  
C;Genetics:  
A;Gene: Tcrz  
C;Keywords: phosphoprotein; T-cell receptor; transmembrane protein

Query Match 25.9%; Score 609; DB 2; Length 164;  
Best Local Similarity 80.7%; Pred. No. 6.5e-33;  
Matches 121; Conservative 7; Mismatches 18; Indels 4; Gaps 2;

QY 295 GGAHTRGDLFDALPKCYLLDGLIFLYGVLTALFLRVKFSRSAPPAQQGQNLN 354  
DB 18 GAAQSGFL---LDPKLCYLLDGLIFLYGVLTALFLRAKFSRSAPPAQQGQNLN 74  
QY 355 LNLGRREYDVLDRGRDPEMGK-PRKNPQGLYNELOKDKMAEAYSEIGMKGERR 413  
DB 75 LNLGRREYDVLDRGRDPEMGKQRRNPQGVYALQDKMAEAYSEIGTKGERR 134  
QY 414 KGHDGLYQGLSTATKDTYDALHMQALPPR 443  
DB 135 KGHDGLYQGLSTATKDTYDALHMQTLAPR 164

RESULT 3  
JC4664  
T-cell receptor zeta chain - bovine  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 10-May-1996 #sequence\_revision 19-Jul-1996 #text\_change 05-Nov-1999  
A;Accession: JC4664  
R;Hagens, G.; Galley, Y.; Glaser, I.; Davis, W.C.; Baldwin, C.L.; Clevers, H.; Dobbela  
Gene 169, 165-171, 1996  
A;Title: Cloning, sequencing and expression of the bovine CD3 epsilon and TCR-zeta chain  
A;Reference number: JC4663; MUID:96194796; PMID:8647441  
A;Accession: JC4664  
A;Molecule type: mRNA  
A;Residues: 1-166. <HAG>  
A;Cross-references: GB:U25688; NID:g1263011; PIDN:AAC48548.1; PID:g1263012  
C;Comment: This protein plays a pivotal role in linking T-cell receptor-triggering to s  
ymphokine receptor gene expression.  
C;Genetics:  
A;Gene: tcr-zeta  
C;Keywords: GTP binding; signal transduction; T-cell receptor  
F:129-146/Region::GDP/GTP-binding

Query Match 25.7%; Score 605; DB 2; Length 166;  
Best Local Similarity 80.7%; Pred. No. 1.2e-32;  
Matches 121; Conservative 9; Mismatches 14; Indels 6; Gaps 3;

QY 297 AVHTRGDLFDALPKCYLLDGLIFLYGVLTALFLRVKFSRSAPPAQQGQNLN 356  
DB 20 AAQSGFL---LDPKLCYLLDGLIFLYGVLTALFLRAKFSRSANAPAYQQGQNPVYN 76  
QY 357 LGRREYDVLDRGRDPEMGK-PRKNPQGLYNELOKDKMAEAYSEIGMK--GERR 413  
DB 77 VGRREYAVLDRGGFDPEMGKQRRNPQGVYALQDKMAEAYSEIGMKSDNQR 136  
QY 414 KGHDGLYQGLSTATKDTYDALHMQALPPR 443  
DB 137 KGHDGLYQGLSTATKDTYDALHMQALPPR 166

RESULT 4  
JC4624  
T-cell surface glycoprotein CD3 zeta chain - sheep  
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C;Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 09-Jul-2004  
A;Accession: I46424; S22980  
R;Hein, W.R.; Tunnacliffe, A.

Immunogenetics 37, 279-284, 1993  
A;Title: Invariant components of the sheep T-cell antigen receptor: cloning of the CD3 e  
A;Reference number: I46424; MUID:93131305; PMID:8420837  
A;Accession: I46424  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-166 <HEI>  
A;Cross-references: UNIPROT:P29329; EMBL:Z12968; NID:g1399; PIDN:CAA78312.1; PID:g1400  
C;Keywords: glycoprotein

Query Match 25.7%; Score 605; DB 2; Length 166;  
Best Local Similarity 80.0%; Pred. No. 1.2e-32;  
Matches 120; Conservative 11; Mismatches 13; Indels 6; Gaps 3;

QY 297 AVHTRGDLFDALPKCYLLDGLIFLYGVLTALFLRVKFSRSAPPAQQGQNLN 356  
DB 20 AAQSGFL---LDPKLCYLLDGLIFLYGVLTALFLRAKFSRSADAPAYQHGQNPVYN 76  
QY 357 LGRREYDVLDRGRDPEMGK-PRKNPQGLYNELOKDKMAEAYSEIGMK--GERR 413  
DB 77 VGRREYAVLDRGGFDPEMGKQRRNPQGVYALQDKMAEAYSEIGMKSDNQR 136  
QY 414 KGHDGLYQGLSTATKDTYDALHMQALPPR 443  
DB 137 KGHDGLYQGLSTATKDTYDALHMQALPPR 166

RESULT 5  
A45089  
T-cell receptor CD3 theta chain, alternate splice form - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 30-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 05-Nov-1999  
A;Accession: A45089; I49587  
R;Clayton, L.K.; Diener, A.C.; Lerner, A.; Tse, A.G.; Koyasu, S.; Reinherz, E.L.  
J. Biol. Chem. 267, 26023-26030, 1992  
A;Title: Differential regulation of T-cell receptor processing and surface expression af  
A;Reference number: A45089; MUID:93100325; PMID:1464613  
A;Accession: A45089  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-188 <CIA>  
A;Cross-references: GB:S51932; NID:g261998; PIDN:AAB24559.1; PID:g261999  
A;Experimental source: thymus  
A;Note: sequence extracted from NCBI backbone (NCBIP:120865)  
R;Lerner, A.; Diener, A.C.; Reinherz, E.L.; Clayton, L.K.  
Eur. J. Immunol. 22, 2135-2140, 1992  
A;Title: Human genomic sequences corresponding to murine CD3eta-related transcripts: lac  
A;Reference number: I49587; MUID:92347411; PMID:1322304  
A;Accession: I49587  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-188 <RES>  
A;Cross-references: GB:L03353; NID:g192508; PIDN:AAA37401.1; PID:g192509  
C;Keywords: T-cell receptor

Query Match 22.2%; Score 521.5; DB 2; Length 188;  
Best Local Similarity 79.7%; Pred. No. 3.8e-27;  
Matches 102; Conservative 7; Mismatches 16; Indels 3; Gaps 1;

QY 295 GGAHTRGDLFDALPKCYLLDGLIFLYGVLTALFLRVKFSRSAPPAQQGQNLN 354  
DB 18 GAAQSGFL---LDPKLCYLLDGLIFLYGVLTALFLRAKFSRSANAPAYQQGQNPVYN 74  
QY 355 LNLGRREYDVLDRGRDPEMGKQRRNPQGVYALQDKMAEAYSEIGMKGERR 414  
DB 75 LNLGRREYDVLDRGRDPEMGKQRRNPQGVYALQDKMAEAYSEIGTKGERR 134  
QY 415 KGHDGLYQ 422  
DB 135 KGHDGLYQ 142

RESULT 6

A;Note: this chain was isolated from an IgM myeloma protein that binds gamma-  
E;Miller III, A.; Glaser, J.A.  
J. Mol. Biol. 205, 763-778, 1989  
A;Title: Comparative sequence and immunochemical analyses of murine monoclonal anti-morpho-  
A;Reference number: S06815; MUID:90064531; PMID:2555519  
A;Accession: S06815  
A;Status: not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-118 <MIL>



A:Experimental source: clone IIIB  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:15-98/Domain: immunoglobulin homology <IMM>  
F:22-96/Disulfide bonds: #status predicted

Query Match 20.8%; Score 486.5; DB 1; Length 119;  
Best Local Similarity 75.8%; Pred. No. 3.2e-25;  
Matches 91; Conservative 15; Mismatches 11; Indels 3; Gaps 2;  
  
QY 142 EVQLVESGGGVQPGSRSLRLSCSASGDFPTTYWMSWVRQAPGKLEWIGEIHDPDSSTINY 201  
DB 1 EVKLESGGGLVQPGGSLKLSCAASGDFSRFYWMSWVRQAPGKLEWIGEIHDPDSSTINY 60  
  
QY 202 APSLKDRFTISRDNKNTLFLQMSLRPEDTGVPFCASL-YFGFPWFAYWGQTPVTYSS 260  
DB 61 TPLSKDKFTISRDNKNTLYLQMSKVRSEDTALYCARLHYGYA--AYWGQGLTVTVSA 118

RESULT 10  
AVMSJ5  
Ig heavy chain V region (J539) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Apr-1980 #sequence\_revision 30-Apr-1980 #text\_change 09-Jul-2004  
C:Accession: A02080  
R:Rao, D.N.; Rudikoff, S.; Krutzsch, H.; Potter, M.  
Proc. Natl. Acad. Sci. U.S.A. 76, 2890-2894, 1979  
A:Title: Structural evidence for independent joining region gene in immunoglobulin heavy  
etermining regions.  
A:Reference number: A93832; MUID:79223895; PMID:111245  
A:Accession: A02080  
A:Molecule type: protein  
A:Residues: 1-119 <RAO>  
A:Cross-references: UNIPROT:P01810  
C:Comment: This chain was isolated from a myeloma protein that binds galactan.  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:15-98/Domain: immunoglobulin homology <IMM>  
F:22-96/Disulfide bonds: #status predicted

Query Match 20.7%; Score 486.5; DB 1; Length 119;  
Best Local Similarity 75.0%; Pred. No. 4.3e-25;  
Matches 90; Conservative 15; Mismatches 12; Indels 3; Gaps 2;  
  
QY 142 EVQLVESGGGVQPGSRSLRLSCSASGDFPTTYWMSWVRQAPGKLEWIGEIHDPDSSTINY 201  
DB 1 EVKLESGGGLVQPGGSLKLSCAASGDFSRFYWMSWVRQAPGKLEWIGEIHDPDSSTINY 60  
  
QY 202 APSLKDRFTISRDNKNTLFLQMSLRPEDTGVPFCASL-YFGFPWFAYWGQTPVTYSS 260  
DB 61 TPLSKDKFTISRDNKNTLYLQMSKVRSEDTALYCARLHYGYN--AYWGQGLTVTVSA 118

RESULT 11  
A25912  
Ig heavy chain precursor V region (W3129) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 23-Jul-1999  
C:Accession: A25912  
R:Borden, P.; Kabat, E.A.  
Proc. Natl. Acad. Sci. U.S.A. 84, 2440-2443, 1987  
A:Title: Nucleotide sequence of the cDNAs encoding the variable region heavy and light c  
etermining regions.  
A:Reference number: A94147; MUID:87175689; PMID:2436230  
A:Accession: A25912  
A:Molecule type: mRNA  
A:Residues: 1-139 <BOR>  
A:Cross-references: GB:M15873; NID:g195268; PIDN:AAA38228.1; PID:g195269  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-139/Product: Ig heavy chain V region W3129 #status predicted <VAR>  
F:33-116/Domain: immunoglobulin homology <IMM>

Query Match 20.5%; Score 482; DB 2; Length 139;  
Best Local Similarity 72.0%; Pred. No. 1e-24;  
Matches 90; Conservative 15; Mismatches 18; Indels 2; Gaps 1;  
  
QY 138 GSGSEVOLVESGGGVQPGSRSLRLSCSASGDFPTTYWMSWVRQAPGKLEWIGEIHDPDS 197  
DB 15 GVQCEVKVIESGGGLVQPGGSLKLSCAASGDFSRFYWMSWVRQAPGKLEWIGEINPDSS 74  
  
QY 198 TINTAPSLKDRFTISRDNKNTLFLQMSLRPEDTGVPFCASL-YFGFPWFPA--YWGQGTTP 255  
DB 75 TINTPSSLKDRFTISRDNKNTLYLQMSKVRSEDTALYCARLGGDLHYAMDYWGQGT 134  
  
QY 256 VTVSS 260  
DB 135 VTVSS 139

RESULT 12  
S06817  
Ig heavy chain V region (clone 11C7) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Sep-1991 #sequence\_revision 11-Nov-1994 #text\_change 23-Jul-1999  
C:Accession: S06817  
R:Miller III, A.; Glaeser, J.A.  
J. Mol. Biol. 209, 763-778, 1989  
A:Title: Comparative sequence and immunochemical analyses of murine monoclonal anti-morp  
A:Reference number: S06815; MUID:90064531; PMID:2555519  
A:Accession: S06817  
A:Molecule type: mRNA  
A:Residues: 1-130 <MIL>  
A:Cross-references: EMBL:X17165; NID:g51917; PIDN:CAA35043.1; PID:g930155  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:15-98/Domain: immunoglobulin homology <IMM>  
F:22-96/Disulfide bonds: #status predicted

Query Match 20.5%; Score 481; DB 2; Length 130;  
Best Local Similarity 70.5%; Pred. No. 1.1e-24;  
Matches 93; Conservative 18; Mismatches 13; Indels 8; Gaps 4;  
  
QY 142 EVQLVESGGGVQPGSRSLRLSCSASGDFPTTYWMSWVRQAPGKLEWIGEIHDPDSSTINY 201  
DB 1 EVNLES GGGLVQPGGSLKLSCAASGDFSRFYWMSWVRQAPGKLEWIGEINPDSSNTINY 60  
  
QY 202 APSLKDRFTISRDNKNTLFLQMSLRPEDTGVPFCA---SLY-FG--FPMFAYWGQGTTP 255  
DB 61 TPLSKDRFTISRDNKNTLYLQMSKVRPEDTGLYCAWTKSVYNGSHYYFDVWGAGTT 120  
  
QY 256 VTVSSAKPTTTP 267  
DB 121 VTVSSAK--TTP 130

RESULT 13  
AVMST6  
Ig heavy chain V region (T601) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Apr-1980 #sequence\_revision 30-Apr-1980 #text\_change 09-Jul-2004  
C:Accession: A02078  
R:Rao, D.N.; Rudikoff, S.; Krutzsch, H.; Potter, M.  
Proc. Natl. Acad. Sci. U.S.A. 76, 2890-2894, 1979  
A:Title: Structural evidence for independent joining region gene in immunoglobulin heavy  
etermining regions.  
A:Reference number: A93832; MUID:79223895; PMID:111245  
A:Accession: A02078  
A:Molecule type: protein  
A:Residues: 1-119 <RAO>  
A:Cross-references: UNIPROT:P01808  
C:Comment: This chain was isolated from an IgA myeloma protein that binds galactan.  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:15-98/Domain: immunoglobulin homology <IMM>  
F:22-96/Disulfide bonds: #status predicted

Query Match 20.4%; Score 480.5; DB 1; Length 119;  
Best Local Similarity 75.0%; Pred. No. 1.1e-24;  
Matches 90; Conservative 14; Mismatches 13; Indels 3; Gaps 2;

QY 142 EVOLVESGGVOPGSRSLRLSCASGDFDTTYMWSVVRQAPGKLEWIGIHPDSS 201  
DB 1 EVKLLSEGGGLVOPGGSLKUSCAASGDFSRYYMWSVVRQAPGKLEWIGIHPDSS 60  
QY 202 APSLKDRFTISRDNKNTLFLQMSLRPEDTGVYFCASL-YFGFPWFAYWQGTPTVSS 260  
DB 61 TPSLKDKFIISRDNAKNTLYLQMSKVRSEDYALYYCARLGYGY--FDVWGAGTTVTSS 118

## RESULT 14

A47329  
IG heavy chain V region (PRI) - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 21-Sep-1993 #sequence\_revision 17-Jul-1994 #text\_change 30-May-1997  
C;Accession: A47329  
R;Brinkmann, U.; Gallo, M.; Brinkmann, E.; Kunwar, S.; Pastan, I.  
Proc. Natl. Acad. Sci. U.S.A. 90, 547-551, 1993  
A;Title: A recombinant immunotoxin that is active on prostate cancer cells and that is  
A;Reference number: A47329; MUID:93133825; PMID:8421689  
A;Accession: A47329  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-118 <BRI>  
A;Experimental source: Balb/c  
A;Note: sequence modified after extraction from NCBI backbone  
A;Note: sequence extracted from NCBI backbone (NCBIN:122874)  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin  
F;16-99/Domain: immunoglobulin homology <IMM>

Query Match 20.4%; Score 475; DB 2; Length 118;  
Best Local Similarity 74.8%; Pred. No. 2.4e-24;  
Matches 89; Conservative 12; Mismatches 16; Indels 2; Gaps 1;

QY 142 EVOLVESGGVOPGSRSLRLSCASGDFDTTYMWSVVRQAPGKLEWIGIHPDSS 201  
DB 2 DVQLVESGGGLVOPGGSLKUSCAASGDFSRYYMWSVVRQAPGKLEWIGIHPDSS 61  
QY 202 APSLKDRFTISRDNKNTLFLQMSLRPEDTGVYFCASL-YFGFPWFAYWQGTPTVSS 260  
DB 62 TPSLKDKFIISRDNAKNTLYLQMSKVRSEDYALYYCARR--GYAMDYWGQGTSTVSS 118

## RESULT 15

S13791  
IG heavy chain V region (X-24) - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 18-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 16-Aug-1996  
C;Accession: S13791  
R;Arepalli, S.R.; Heller, M.; Glaudemans, C.P.J.  
Nucleic Acids Res. 18, 7152, 1990  
A;Title: Sequence of the V(H) gene for murine IgA X-24.  
A;Reference number: S13791; MUID:91088313; PMID:2124679  
A;Accession: S13791  
A;Molecule type: DNA  
A;Residues: 1-136 <ARE>  
A;Cross-references: EMBL:X55984  
A;Note: the authors translated the codon GGG for residue 15 as Lys  
C;Genetics:  
A;Introns: 15/1  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin  
F;33-116/Domain: immunoglobulin homology <IMM>

Query Match 20.1%; Score 473.5; DB 2; Length 136;  
Best Local Similarity 71.8%; Pred. No. 3.5e-24;  
Matches 89; Conservative 14; Mismatches 18; Indels 3; Gaps 2;

QY 138 GSGSEVOLVESGGVOPGSRSLRLSCASGDFDTTYMWSVVRQAPGKLEWIGIHPDSS 197  
DB 15 GVOCEYKLLSEGGGLVOPGGSLNLSCAASGDFSRYYMWSVVRQAPGKLEWIGIHPDSS 74  
QY 198 TINYAPSLKDRFTISRDNKNTLFLQMSLRPEDTGVYFCASL-YFGFPWFAYWQGTPTV 256  
DB 75 TINYTESLKDKFIISRDNAKNTLYLQMSKVRSEDYALYYCARLGYGY--FDVWGQGTTL 132  
QY 257 TVSS 260  
DB 133 TVSS 136

Search completed: May 26, 2005, 02:15:06  
Job time : 69 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2005, 19:58:54 ; Search time 345 Seconds  
(without alignments)  
657.539 Million cell updates/sec

Title: US-10-006-771B-2

Perfect score: 2350

Sequence: 1 MGWSCIILFLVATGVHSD.....LSTATKDTYDALHMQALPPR 443

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	712	30.3	163	1 CD3Z_PIG	Q9XJ99 sus scrofa
2	701.5	29.9	164	1 CD3Z_HUMAN	P20963 homo sapien
3	626.5	26.7	165	1 CD3Z_RABIT	Q9Tuf8 oryctolagus
4	611	26.0	262	2 Q65ZT1	Q65Z11 mus musculus
5	609	25.9	164	1 CD3Z_MOUSE	P24161 mus musculus
6	605	25.7	166	1 CD3Z_SHEEP	P29329 ovis aries
7	511	21.7	206	1 CD3H_MOUSE	P29020 mus musculus
8	488.5	20.8	119	1 HV37_MOUSE	P01807 mus musculus
9	486.5	20.7	119	1 HV40_MOUSE	P01810 mus musculus
10	480.5	20.4	119	1 HV38_MOUSE	P01808 mus musculus
11	465.5	19.8	298	2 Q9QYF0	Q9QYF0 synthetic c
12	465	19.8	117	1 HV41_MOUSE	P01811 mus musculus
13	462	19.7	118	1 HV39_MOUSE	P01809 mus musculus
14	445.5	19.0	126	1 HV3K_HUMAN	P01772 homo sapien
15	445	18.9	117	1 HV42_MOUSE	P01812 mus musculus
16	443.5	18.9	108	1 KV1M_HUMAN	P01605 homo sapien
17	443	18.9	322	2 Q6KAV0	Q6KAV0 homo sapien
18	442	18.8	493	2 Q6GMX2	Q6GMX2 homo sapien
19	435.5	18.5	470	2 Q6PUA4	Q6PUA4 homo sapien
20	435	18.5	236	2 Q6GMX9	Q6GMX9 homo sapien
21	431.5	18.4	108	1 KV1V_HUMAN	P00362 homo sapien
22	431	18.3	116	1 HV36_MOUSE	P01806 mus musculus
23	429.5	18.3	108	1 KV1B_HUMAN	P01594 homo sapien
24	429.5	18.3	108	2 Q9UL70	Q9UL70 homo sapien
25	429.5	18.3	478	2 Q6PI81	Q6PI81 homo sapien
26	429	18.3	147	2 Q9Y509	Q9Y509 homo sapien
27	428.5	18.2	613	2 Q8WUK1	Q8WUK1 homo sapien
28	428	18.2	236	2 Q7Z3Y4	Q7Z3Y4 homo sapien
29	427.5	18.2	108	1 KV1A_HUMAN	P01593 homo sapien
30	426.5	18.1	108	2 Q9UL77	Q9UL77 homo sapien
31	426.5	18.1	573	2 Q8WU38	Q8WU38 homo sapien

32	426	18.1	107	1 KV1D_HUMAN	P01596 homo sapien
33	426	18.1	236	2 Q6GMW1	Q6GMW1 homo sapien
34	426	18.1	236	2 Q6GMX8	Q6GMX8 homo sapien
35	426	18.1	236	2 Q6PIT5	Q6PIT5 homo sapien
36	425.5	18.1	108	1 KV1O_HUMAN	P01607 homo sapien
37	424.5	18.1	108	1 KV1P_HUMAN	P01608 homo sapien
38	424.5	18.1	108	1 KV1V_HUMAN	P04430 homo sapien
39	423.5	18.0	108	1 KV1H_HUMAN	P01600 homo sapien
40	423	18.0	107	2 Q96SA9	Q96SA9 homo sapien
41	422	18.0	236	2 Q6PIH7	Q6PIH7 homo sapien
42	421	17.9	487	2 Q6ZVK0	Q6ZVK0 homo sapien
43	419.5	17.9	108	1 KV1K_HUMAN	P01603 homo sapien
44	419.5	17.9	485	2 Q6PDB8	Q6PDB8 mus musculus
45	419.5	17.9	606	2 Q6GM12	Q6GM12 homo sapien

## ALIGNMENTS

RESULT 1					
CD3Z_PIG					
ID	CD3Z_PIG	STANDARD;	PRT;	163 AA.	
AC	Q9XJ99;				
DT	28-FEB-2003 (Rel. 41, Created)				
DT	28-FEB-2003 (Rel. 41, Last sequence update)				
DT	25-OCT-2004 (Rel. 45, Last annotation update)				
DE	T-cell surface glycoprotein CD3 zeta chain precursor (T-cell receptor T3 zeta chain).				
DE	T3 zeta chain).				
GN	Names=CD3Z;				
OS	Sus scrofa (Pig).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.				
OX	NCBI_TaxID=9823;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Minnesota miniature swine;				
RA	Jie H.-B., Yim D., Kim Y.B.;				
RT	"The molecular cloning of porcine CD3 zeta.";				
RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.				
CC	-I- FUNCTION: Probable role in assembly and expression of the TCR complex as well as signal transduction upon antigen triggering.				
CC	-I- SUBUNIT: The TCR/CD3 complex of T lymphocytes consists of either a TCR alpha/beta or TCR gamma/delta heterodimer coexpressed at the cell surface with the invariant subunits of CD3 labeled gamma, delta, epsilon, zeta, and eta (by similarity).				
CC	-I- SUBCELLULAR LOCATION: Type I membrane protein.				
CC	-I- PTM: Phosphorylated on Tyr residues after T-cell receptor triggering (by similarity).				
CC	-I- SIMILARITY: Belongs to the CD3Z/FCER1G family.				
CC	-I- SIMILARITY: Contains 3 ITAM domains.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC	-----				
CC	EMBL; AF153830; AAD34640.1; -				
DR	InterPro; IPR003110; ITAM.				
DR	Pfam; PF02189; ITAM; 3.				
DR	SMART; SM00077; ITAM; 3.				
KW	Phosphorylation; Receptor; Repeat; Signal; T-cell; Transmembrane.				
FT	CHAIN 1 21				
FT	By similarity.				
FT	T-cell surface glycoprotein CD3 zeta chain.				
FT	SIGNAL 22 163				
FT	DOMAIN 22 30				
FT	Extracellular (Potential).				
FT	TRANSMEM 31 51				
FT	Potential.				
FT	DOMAIN 52 163				
FT	Cytoplasmic (Potential).				
FT	DOMAIN 69 89				
FT	ITAM 1.				
FT	DOMAIN 107 128				
FT	ITAM 2.				
FT	DOMAIN 138 158				
FT	ITAM 3.				

```
FT DISULFID 32 32 Interchain (Potential).
RL MOD RES 152 152 Phosphotyrosine (By similarity).
SQ SEQUENCE 163 AA; 18568 MW; 34898620B67167C7 CRC64;

Query Match 30.3%; Score 712; DB 1; Length 163;
Best Local Similarity 98.5%; Pred. No. 5.2e-43;
Matches 135; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 307 LDPKLCYLLDGLFLFYGVILTPURVVFSSAEPAYQGGONQYNLNLGRREYDVL 366
DB 27 LDPKLCYLLDGLFLFYGVILTPURVVFSSADAPAYQGGONQYNLNLGRREYDVL 86
QY 367 DKRRGRDEMGKPKRKNPQEGLYNELQDKMAEYSEIGMKGRRRGKHDGLYQGLST 426
DB 87 DKRRGRDEMGKPKRKNPQEGLYNELQDKMAEYSEIGMKGRRRGKHDGLYQGLST 146
QY 427 ATKDTYDALHMQALPPR 443
DB 147 ATKDTYDALHMQALPPR 163

RESULT 2
CD3Z_HUMAN
ID CD3Z_HUMAN STANDARD; PRT; 164 AA.
AC P20963; Q8TAX4;
DT 01-FEB-1991 (Rel. 17, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE T-cell surface glycoprotein CD3 zeta chain precursor (T-cell receptor
DE T3 zeta chain)
GN Name=CD3Z; Synonyms=T3Z, TCRZ;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=89071765; PubMed=2974162;
RA Weissman A.M., Hou D., Orloff D.G., Modi W.S., Seunanez H.,
RA O'Brien S.J., Klausner R.D.;
RT "Molecular cloning and chromosomal localization of the human T-cell
RT receptor zeta chain: distinction from the molecular CD3 complex.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:9709-9713(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Roak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S.E., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP INTERACTION WITH HIV-1 NEF.
RX PubMed=10224289;
RA Xu X.-N., Laffert B., Screaton G.R., Kraft M., Wolf D., Kolanus W.,
RA Mongkolsapay J., McMichael A.J., Baur A.S.;
RT "Induction of Fas ligand expression by HIV involves the interaction of
```

```
RT Nef with the T cell receptor zeta chain.";
RL J. Exp. Med. 189:1489-1496(1999).
RN [4]
RP INTERACTION WITH SLA.
RX MEDLINE=93380595; PubMed=10449770; DOI=10.1073/pnas.96.17.9775;
RA Tang J., Sawadkosol S., Chang J.-H., Burakoff S.J.;
RT "SLAP, a dimeric adapter protein, plays a functional role in T cell
RT receptor signaling.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:9775-9780(1999).
RN [5]
RP INTERACTION WITH DOCK2.
RX MEDLINE=22165501; PubMed=12176041; DOI=10.1016/S0006-291X(02)00931-2;
RA Nishihara H., Maeda M., Tsuda M., Makino Y., Sawa H., Nagashima K.,
RA Tanaka S.;
RT "DOCK2 mediates T cell receptor-induced activation of Rac2 and IL-2
RT transcription.";
RL Biochem. Biophys. Res. Commun. 296:716-720(2002).
RN [6]
RP PHOSPHORYLATION SITES TYR-83; TYR-111; TYR-123 AND TYR-142.
RX PubMed=12522270; DOI=10.1073/pnas.24361911100;
RA Salomon A.R., Ficarro S.B., Brill L.M., Brinker A., Phung Q.T.,
RA Ericson C., Sauer K., Brock A., Horn D.M., Schultz P.G., Peters E.C.;
RT "Profiling of tyrosine phosphorylation pathways in human cells using
RT mass spectrometry.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:443-448(2003).
RN [7]
RP STRUCTURE BY NMR OF 136-149.
RX MEDLINE=93201600; PubMed=7680960; DOI=10.1016/0092-8674(93)90405-F;
RA Wakeman G., Shoelson S.E., Pant N., Cowburn D., Kuriyan J.;
RT "Binding of a high affinity phosphotyrosyl peptide to the Src SH2
RT domain: crystal structures of the complexed and peptide-free forms.";
RL Cell 72:779-790(1993).
CC -!- FUNCTION: Probable role in assembly and expression of the TCR
CC complex as well as signal transduction upon antigen triggering.
CC -!- SUBUNIT: The TCR/CD3 complex of T lymphocytes consists of either a
CC TCR alpha/beta or TCR gamma/delta heterodimer coexpressed at the
CC cell surface with the invariant subunits of CD3 labeled gamma,
CC delta, epsilon, zeta, and eta. CD3-zeta forms either homodimers or
CC heterodimers with CD3-eta. Interacts with SLA and SLA2. Interacts
CC with DOCK2. Interacts with HIV-1 Nef protein.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=CD-3-zeta;
CC IsoId=P20963-1; Sequence=Displayed;
CC Name=CD-3-eta;
CC IsoId=P20963-2; Sequence=Not described;
CC -!- PTM: Phosphorylated on Tyr residues after T-cell receptor
CC triggering (By similarity).
CC -!- SIMILARITY: Belongs to the CD3Z/FCER1G family.
CC -!- SIMILARITY: Contains 3 ITAM domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J04132; AAA60394.1; -.
CC EMBL; BC025703; AAH25703.1; -.
CC PIR; A31768; A31768.
CC PDB; 1TCE; NMR; B=136-149.
CC Genew; HGNC:1677; CD3Z.
CC H-invDB; HIX0001296; -.
CC MIM; 186780; -.
CC GO; GO:0005886; C:plasma membrane; TAS.
CC GO; GO:0042101; C:l-cell receptor complex; TAS.
CC GO; GO:0005515; F:protein binding; IPI.
CC GO; GO:0042803; F:protein homodimerization activity; NAS.
CC InterPro; IPR003110; ITAM.
CC Pfam; PF02189; ITAM; 3.
```

```

DR SMART; SM00077; ITAM; 3.
KW 3D-structure; Alternative splicing; Phosphorylation; Receptor; Repeat;
KW SIGNAL; T-cell; Transmembrane.
FT SIGNAL 1 21
FT CHAIN 22 164
FT T-cell surface glycoprotein CD3 zeta
FT chain.
FT DOMAIN 22 30 Extracellular (Potential).
FT TRANSMEM 31 51 Potential.
FT DOMAIN 52 164 Cytoplasmic (Potential).
FT DOMAIN 69 89 ITAM 1.
FT DOMAIN 108 129 ITAM 2.
FT DOMAIN 139 159 ITAM 3.
FT DISULFID 32 32 Interchain (Potential).
FT MOD RES 83 83 Phosphotyrosine.
FT MOD RES 111 111 Phosphotyrosine.
FT MOD RES 123 123 Phosphotyrosine.
FT MOD RES 142 142 Phosphotyrosine.
FT MOD RES 153 153 Phosphotyrosine.
FT MOD RES 161 161 DA -> EP (in Ref. 1).
FT CONFLICT 101 101 Missing (in Ref. 1).
FT CONFLICT 101 101 Missing (in Ref. 1).
SQ SEQUENCE 164 AA; 18696 MW; 9408260374856EE9 CRC64;

Query Match 29.9%; Score 701.5; DB 1; Length 164;
Best Local Similarity 97.8%; Pred. No. 2.9e-42;
Matches 135; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

Qy 307 LDPKLCYLLDGLIYGVILTAFLRVKFSRAEPAYQQQNOLYNELNLRREYDVL 366
Db 27 LDPKLCYLLDGLIYGVILTAFLRVKFSRAEPAYQQQNOLYNELNLRREYDVL 86
Qy 367 DKRRGRDPENGGKP-RRKNPQEGLYNELQDKMAEAYSEIGMKERRRGKHDGLYGLS 425
Db 87 DKRRGRDPENGGKPQRRKNPQEGLYNELQDKMAEAYSEIGMKERRRGKHDGLYGLS 146
Qy 426 TATKDTYDALHMQALPPR 443
Db 147 TATKDTYDALHMQALPPR 164

RESULT 3
CD3Z_RABIT STANDARD; PRT; 165 AA.
ID Q65Z11 PRELIMINARY; PRT; 262 AA.
AC Q65Z11 PRELIMINARY; PRT; 262 AA.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE T-cell surface glycoprotein CD3 zeta chain precursor (T-cell receptor
DE T3 zeta chain).
GN Name=CD3Z;
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B/J x Chbb:HM;
RA Isono T, Nishimura M.;
RT "Rabbit CD3 zeta."
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Probable role in assembly and expression of the TCR
CC complex as well as signal transduction upon antigen triggering.
CC -!- SUBUNIT: The TCR/CD3 complex of T lymphocytes consists of either a
CC TCR alpha/beta or TCR gamma/delta heterodimer coexpressed at the
CC cell surface with the invariant subunits of CD3 labeled gamma,
CC delta, epsilon, zeta, and eta (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- PTM: Phosphorylated on Tyr residues after T-cell receptor
CC triggering (By similarity).
CC -!- SIMILARITY: Belongs to the CD3Z/FCER1G family.
CC -!- SIMILARITY: Contains 3 ITAM domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB035152; BAA86994.1; -.
DR InterPro; IPR003110; ITAM.
DR Pfam; PF02189; ITAM; 3.
DR SMART; SM00077; ITAM; 3.
KW Phosphorylation; Receptor; Repeat; Signal; T-cell; Transmembrane.
FT SIGNAL 1 21
FT CHAIN 22 165
FT T-cell surface glycoprotein CD3 zeta
FT chain.
FT DOMAIN 22 30 Extracellular (Potential).
FT TRANSMEM 31 51 Potential.
FT DOMAIN 52 165 Cytoplasmic (Potential).
FT DOMAIN 69 89 ITAM 1.
FT DOMAIN 107 128 ITAM 2.
FT DOMAIN 140 160 ITAM 3.
FT DISULFID 32 32 Interchain (Potential).
FT MOD RES 154 154 Phosphotyrosine (By similarity).
SQ SEQUENCE 165 AA; 18773 MW; 3183136130BAA4F5 CRC64;

Query Match 26.7%; Score 626.5; DB 1; Length 165;
Best Local Similarity 83.0%; Pred. No. 6.5e-37;
Matches 122; Conservative 5; Mismatches 15; Indels 5; Gaps 2;

Qy 299 HTRGLDFALDPKLCYLLDGLIYGVILTAFLRVKFSRAEPAYQQQNOLYNELNIG 358
Db 22 HIFGL---LDPKLCYLLDGLIYGVILTAFLRVKFSRAEPAYQQQNOLYNELNIG 78
Qy 359 RREYDVLDRGRDPEMGKPKRRKNPQEGLYNELQDKMAEAYSEIGMKGE--RRRGKG 416
Db 79 RREYDVLDRGRDPEMGKPKRRKNPQEGLYNELQDKMAEAYSEIGMKGENQRREGK 138
Qy 417 HDGLYQGLSTATKDTYDALHMQALPPR 443
Db 139 HDGLYQGLSNAATKDTYDALHMQALPPR 165

RESULT 4
Q65Z11 PRELIMINARY; PRT; 262 AA.
ID Q65Z11 PRELIMINARY; PRT; 262 AA.
AC Q65Z11 PRELIMINARY; PRT; 262 AA.
DT 25-OCT-2004 (T-EMBLrel. 28, Created)
DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
DE Anti-HIV-1 reverse transcriptase single-chain variable.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hybridoma;
RX MEDLINE=96211469; PubMed=8648670;
RA Shaheen F., Duan L., Zhu M., Bagasra O., Pomerantz R.J.;
RT "Targeting human immunodeficiency virus type 1 reverse transcriptase
RT by intracellular expression of single-chain variable fragments to
RT inhibit early stages of the viral life cycle."
RL J. Virol. 70:3392-3400(1996).
DR EMBL; U48716; AB64342.1; -.
DR GO; GO:0003364; F:RNA-directed DNA polymerase activity; IEA.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR003596; IG_v.
DR Pfam; PF00047; IG; 2.
DR SMART; SM00409; IG; 2.
DR SMART; SM00408; IGc2; 2.
DR SMART; SM00406; IGv; 2.
DR PROSITE; PS00835; IG_LIKE; 2.

```





```
RT protein-2 (SLAP-2), which negatively regulates T cell receptor
RL J. Biol. Chem. 277:19131-19138(2002).
CC -1- FUNCTION: Probable role in assembly and expression of the TCR
CC complex as well as signal transduction upon antigen triggering.
CC -1- SUBUNIT: Interacts with DOK2 (By similarity). The TCR/CD3 complex
CC of T lymphocytes consists of either a TCR alpha/beta or TCR
CC gamma/delta heterodimer coexpressed at the cell surface with the
CC invariant subunits of CD3 labeled gamma, delta, epsilon, zeta, and
CC eta. CD3-zeta forms either homodimers or heterodimers with CD3-
CC eta. Interacts with SLA and SLA2.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=CD-3-zeta;
CC IsoId=P24161-1; Sequence=Displayed;
CC Name=CD-3-eta;
CC IsoId=P29020-1; Sequence=External;
CC -1- PTM: Phosphorylated on Tyr residues after T-cell receptor
CC triggering.
CC -1- SIMILARITY: Belongs to the CD3Z/FCER1G family.
CC -1- SIMILARITY: Contains 3 ITAM domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M19729; AAA40171.1; -.
DR EMBL; J04967; AAA50301.1; -.
DR EMBL; AK017904; BAB30957.1; -.
DR EMBL; BC052824; AAH52824.1; -.
DR PIR; A40104; A40104.
DR MGD; MGI:88334; CD3z.
DR InterPro; IPR003110; ITAM.
DR Pfam; PF02189; ITAM; 3.
DR Alternative splicing; Direct protein sequencing; Phosphorylation;
KW Receptor; Repeat; Signal; T-cell; Transmembrane.
FT SIGNAL 1 21
FT CHAIN 22 164 T-cell surface glycoprotein CD3 zeta
FT chain.
FT DOMAIN 22 30 Extracellular (Potential).
FT TRANSMEM 31 51 Potential.
FT DOMAIN 52 164 Cytoplasmic (Potential).
FT DOMAIN 69 89 ITAM 1.
FT DOMAIN 108 129 ITAM 2.
FT DOMAIN 139 159 ITAM 3.
FT DISULFID 32 32 Interchain (Potential).
FT MOD_RES 153 153 Phosphotyrosine.
FT CONFLICT 153 153 Y -> C (in Ref. 3).
SQ SEQUENCE 164 AA; 18637 MW; 1B8022035A312831 CRC64;

Query Match 25.9%; Score 609; DB 1; Length 164;
Best Local Similarity 80.7%; Pred. No. 1-le-35;
Matches 121; Conservative 7; Mismatches 18; Indels 4; Gaps 2;

QY 295 GGAHVHTRGLDFALDPKLCYLLDGLIFLYGVILTAFLRVKFSRSAPPPAYCQGNOLYNE 354
DB 18 GAQAQSFGL---LDPKLCYLLDGLIFLYGVILTAFLRVKFSRSAPPPAYCQGNOLYNE 74
QY 355 LNLGRREYVDLKRGRDPMGK-PRKNPQGLYNELOKDKMAEAYSEIGMKGRRR 413
DB 75 LNLGRREYVDLEKRRADPEMGKQKQRRRNPNQEGVYNALQDKMAEAYSEIGTKGRRR 134
QY 414 GKGDHGYQGLSTATKDTYDALHMQALPPR 443
DB 135 GKGDHGYQGLSTATKDTYDALHMQTLAPR 164

RESULT 6
```

```
CD3Z SHEEP STANDARD; PRT; 166 AA.
ID P29329;
AC 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE T-cell surface glycoprotein CD3 zeta chain precursor (T-cell receptor
DE T3 zeta chain).
GN Name=CD3Z;
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Caprinae; Ovis.
OC NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White alpine;
RX MEDLINE=93131305; PubMed=8420837;
RA Hein W.R., Tunncliffe A.;
RT "Invariant components of the sheep T-cell antigen receptor: cloning of
RT the CD3 epsilon and Tcr zeta chains.";
RL Immunogenetics 37:279-284(1993).
CC -1- FUNCTION: Probable role in assembly and expression of the TCR
CC complex as well as signal transduction upon antigen triggering.
CC -1- SUBUNIT: The TCR/CD3 complex of T lymphocytes consists of either a
CC TCR alpha/beta or TCR gamma/delta heterodimer coexpressed at the
CC cell surface with the invariant subunits of CD3 labeled gamma,
CC delta, epsilon, zeta, and eta.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- PTM: Phosphorylated on Tyr residues after T-cell receptor
CC triggering (By similarity).
CC -1- SIMILARITY: Belongs to the CD3Z/FCER1G family.
CC -1- SIMILARITY: Contains 3 ITAM domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z12968; CAA78312.1; -.
DR PIR; I46424; I46424.
DR InterPro; IPR003110; ITAM.
DR Pfam; PF02189; ITAM; 3.
DR SMART; SM00077; ITAM; 3.
KW Phosphorylation; Receptor; Repeat; Signal; T-cell; Transmembrane.
FT SIGNAL 1 21 By similarity.
FT CHAIN 22 166 T-cell surface glycoprotein CD3 zeta
FT chain.
FT DOMAIN 22 30 Extracellular (Potential).
FT TRANSMEM 31 51 Potential.
FT DOMAIN 52 166 Cytoplasmic (Potential).
FT DOMAIN 69 89 ITAM 1.
FT DOMAIN 108 129 ITAM 2.
FT DOMAIN 141 161 ITAM 3.
FT DISULFID 32 32 Interchain (Potential).
FT MOD_RES 155 155 Phosphotyrosine (By similarity).
SQ SEQUENCE 166 AA; 18704 MW; E7DB9AD84E58311A CRC64;

Query Match 25.7%; Score 605; DB 1; Length 166;
Best Local Similarity 80.0%; Pred. No. 2-2e-35;
Matches 120; Conservative 11; Mismatches 13; Indels 6; Gaps 3;

QY 297 AVHTRGLDFALDPKLCYLLDGLIFLYGVILTAFLRVKFSRSAPPPAYCQGNOLYNE 356
DB 20 AAQSFGL---LDPKLCYLLDGLIFLYGVILTAFLRVKFSRSADAPAYQHQNPNVYNELN 76
QY 357 LGRREYVDLKRGRDPMGKPR-RRKNPQGLYNELOKDKMAEAYSEIGMK--GERRR 413
DB 77 VGRREYAVLDRRGDFDPENGKPRKKNPHEVYNELRKDKMAEAYSEIGMKSDNQRR 136
```

```

QY 414 GKGDGLYQGLSTATKDYDALHMQALPPR 443
Db 137 GKGDGLYQGLSTATKDYDALHMQALPPR 166

RESULT 7
CD3H MOUSE
ID CD3H MOUSE STANDARD; PRT; 206 AA.
AC P29020;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE T-cell surface glycoprotein CD3 eta chain precursor (T-cell receptor
DE T3 eta chain).
GN Name=CD3z; Synonyms=CD3h;
OS Mus musculus. (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP MEDLINE=90239005; PubMed=2139725;
RA Jin Y.J., Clayton L.K., Howard F.D., Koyasu S., Sieh M.,
RA Clayton L.K., D'Adamo L., Sieh M., Hussey R.E., Koyasu S.,
RA Reinherz E.L., Howard F.B.;
RT "CD3 eta and CD3 zeta are alternatively spliced products of a common
RT genetic locus and are transcriptionally and/or post-transcriptionally
RT regulated during T-cell development.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:3319-3323 (1990).
RN [2]
RP SEQUENCE OF 144-206 FROM N.A.
RX MEDLINE=91271358; PubMed=1828894;
RA Clayton L.K., D'Adamo L., Sieh M., Hussey R.E., Koyasu S.,
RA Reinherz E.L., Howard F.B.;
RT "CD3 eta and CD3 zeta are alternatively spliced products of a common
RT genetic locus and are transcriptionally and/or post-transcriptionally
RT regulated during T-cell development.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:5202-5206 (1991).
RN [3]
RP SEQUENCE OF 144-206 FROM N.A.
RX PubMed=2150596;
RA Ohno H., Saito T.;
RT "CD3 zeta and eta chains are produced by alternative splicing from a
RT common gene.";
RL Int. Immunol. 2:1117-1119 (1990).
RN [4]
RP ERRATUM.
RA Ohno H., Saito T.;
RL Int. Immunol. 4:1339-1339 (1992).
CC -!- FUNCTION: Probable role in assembly and expression of the TCR
CC complex as well as signal transduction upon antigen triggering.
CC -!- SUBUNIT: The TCR/CD3 complex of T lymphocytes consists of either a
CC TCR alpha/beta or TCR gamma/delta heterodimer coexpressed at the
CC cell surface with the invariant subunits of CD3 labeled gamma,
CC delta, epsilon, zeta, and eta. CD3-eta can be complexed in a
CC heterodimeric form with CD3-zeta subunit. CD3-eta homodimer has
CC not been observed.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=CD-3-eta;
CC IsoId=P29020-1; Sequence=Displayed;
CC Name=CD-3-zeta;
CC IsoId=P24161-1; Sequence=External;
CC -!- SIMILARITY: Belongs to the CD3Z/FCERIG family.
CC -!- SIMILARITY: Contains 3 ITAM domains.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----

```

```

DR EMBL; M33158; AAA37398.1; -.
DR EMBL; M76711; AAA40403.1; -.
DR PIR; A35900; A35900.
DR MGI; 88334; Cd3z.
DR InterPro; IPR003110; ITAM.
DR Pfam; PF02189; ITAM; 2.
DR SMART; SM00077; ITAM; 2.
KW Alternative splicing; Direct protein sequencing; Receptor; Repeat;
KW Signal; T-cell; Transmembrane.
FT SIGNAL 1 21
FT CHAIN 22 206
FT DOMAIN 22 30 T-cell surface glycoprotein CD3 eta
FT TRANSMEM 31 51 chain.
FT DOMAIN 52 206 Extracellular (Potential).
FT DOMAIN 69 89 Potential.
FT DOMAIN 108 129 Cytoplasmic (Potential).
FT DOMAIN 139 159 ITAM 1.
FT DISULFID 32 32 ITAM 2.
FT DISULFID 32 32 ITAM 3.
FT DISULFID 32 32 Interchain (Potential).
SQ SEQUENCE 206 AA; 23339 MW; 829256A2CF44E444 CRC64;

Query Match 21.7%; Score 511; DB 1; Length 206;
Best Local Similarity 79.1%; Pred. No. 1.5e-28;
Matches 102; Conservative 7; Mismatches 16; Indels 4; Gaps 2;

QY 295 GGAVHTRGLDPALDPKLCYLDGILFIYGVILTALFLRVKFSRAEPPAYQGGNOLYNE 354
Db 18 GAQAQSFGL---LDPKLCYLDGILFIYGVITITALLYLRAKFSRAETAANLQDPNOLYNE 74

QY 355 NLGRREYDVLDKRGDRPEMGKG-PRKNPQSGLYNELQDKNAEYSEIGMKGERRR 413
Db 75 NLGRREYDVLEKRAEDPEMGKGQRRRNPQGVYNALQDKVAEYSEIGTKGERRR 134

QY 414 GKGDGLYQ 422
Db 135 GKGDGLYQ 143

RESULT 8
HV37 MOUSE
ID HV37 MOUSE STANDARD; PRT; 119 AA.
AC P01807;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Ig heavy chain V region X44.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=79223895; PubMed=111245;
RA Rao D.N., Rudikoff S., Krutzsch H., Potter M.;
RT "Structural evidence for independent joining region gene in
RT immunoglobulin heavy chains from anti-galactan myeloma proteins and
RT its potential role in generating diversity in complementarity-
RT determining regions.";
RL Proc. Natl. Acad. Sci. U.S.A. 76:2890-2894 (1979).
CC -!- MISCELLANEOUS: This chain was isolated from an IGA myeloma protein
CC that binds galactan.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR PIR; A02077; AVMSX4.
DR HSPF; P01810; 2FBJ.
DR InterPro; IPR007110; Ig-like.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
KW Direct protein sequencing; Immunoglobulin V region.
FT DOMAIN 1 117
FT NON TER 119 119 Ig-like.
SQ SEQUENCE 119 AA; 13246 MW; BC34FC8F31CD41B3 CRC64;

```

```
Query Match      20.8%; Score 488.5; DB 1; Length 119;
Best Local Similarity 75.8%; Pred. No. 3e-27;
Matches 91; Conservative 15; Mismatches 11; Indels 3; Gaps 2;

QY 142 EVLVESGGGVQPGSRSLRSLSCSASGDFDTTYMWSVRQAPGKGLWIGIHPDSSTINY 201
DB 1 EVKLSSGGGLVQPGGSLKLSLSCSASGDFDSRYMWSVRQAPGKGLWIGIHPDSSTINY 60
QY 202 APSLKDRFTISRDNKNTLFLQMSLRPDTGYVFCASL-YFGFPWFAYWGQGTPTVTSS 260
DB 61 TPSLKDKFTISRDNKNTLYLQMSKVRSEDTALYYCARLHYGYA--AYWGQGTTLVTSA 118

RESULT 9
HV40 MOUSE
ID HV40 MOUSE STANDARD; PRT; 119 AA.
AC P01810;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Ig heavy chain V region J539.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP PRELIMINARY SEQUENCE.
RX MEDLINE=79223895; PubMed=111245;
RA Rao D.N., Rudikoff S., Krutzsch H., Potter M.;
RT "Structural evidence for independent joining region gene in
RT immunoglobulin heavy chains from anti-galactan myeloma proteins and
RT its potential role in generating diversity in complementarity-
RT determining regions.";
RL Proc. Natl. Acad. Sci. U.S.A. 76:2890-2894 (1979).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RX MEDLINE=88217852; PubMed=3449853;
RA Suh S.W., Bhat T.N., Navia M.A., Cohen G.H., Rao D.N., Rudikoff S.,
RA Davies D.R.;
RT "The galactan-binding immunoglobulin Fab J539: an X-ray diffraction
RT study at 2.6-A resolution.";
RL Proteins 1:74-80 (1986).
CC -1- MISCELLANEOUS: This chain was isolated from a myeloma protein that
CC binds galactan.
DR PIR; A02080; AVMSJ5.
DR PDB; 2FBJ; X-ray; H-.
DR InterPro; IPR007110; Ig-like.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IG; 1.
DR PROSITE; PS00835; IG LIKE; 1.
KW 3D-structure; Direct protein sequencing; Immunoglobulin V region.
FT NON_TER 119
FT STRAND 3 7
FT STRAND 10 12
FT TURN 14 15
FT STRAND 18 25
FT STRAND 29 31
FT HELIX 34 39
FT STRAND 41 42
FT TURN 45 51
FT STRAND 53 54
FT TURN 58 60
FT TURN 62 67
FT STRAND 68 72
FT STRAND 78 83
FT STRAND 88 90
FT STRAND 92 100
FT TURN 101 103
FT STRAND 104 108
FT STRAND 112 116
FT SEQUENCE 119 AA; 13240 MW; 577B4F1DB675C1F1 CRC64;
```

```
Query Match      20.7%; Score 486.5; DB 1; Length 119;
Best Local Similarity 75.0%; Pred. No. 4.2e-27;
Matches 90; Conservative 15; Mismatches 12; Indels 3; Gaps 2;

QY 142 EVLVESGGGVQPGSRSLRSLSCSASGDFDTTYMWSVRQAPGKGLWIGIHPDSSTINY 201
DB 1 EVKLSSGGGLVQPGGSLKLSLSCSASGDFDSRYMWSVRQAPGKGLWIGIHPDSSTINY 60
QY 202 APSLKDRFTISRDNKNTLFLQMSLRPDTGYVFCASL-YFGFPWFAYWGQGTPTVTSS 260
DB 61 TPSLKDKFTISRDNKNTLYLQMSKVRSEDTALYYCARLHYGYN--AYWGQGTTLVTSA 118

RESULT 10
HV38 MOUSE
ID HV38 MOUSE STANDARD; PRT; 119 AA.
AC P01808;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Ig heavy chain V region T601.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=79223895; PubMed=111245;
RA Rao D.N., Rudikoff S., Krutzsch H., Potter M.;
RT "Structural evidence for independent joining region gene in
RT immunoglobulin heavy chains from anti-galactan myeloma proteins and
RT its potential role in generating diversity in complementarity-
RT determining regions.";
RL Proc. Natl. Acad. Sci. U.S.A. 76:2890-2894 (1979).
CC -1- MISCELLANEOUS: This chain was isolated from an IgA myeloma protein
CC that binds galactan.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR PIR; A02078; AVMS76.
DR HSSP; P01810; 2FBJ.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IG; 1.
DR PROSITE; PS00835; IG LIKE; 1.
KW Direct protein sequencing; Immunoglobulin V region.
FT DOMAIN 1 112
FT NON_TER 119
FT SEQUENCE 119 AA; 13169 MW; BC38CC84E5EA00E8 CRC64;

Query Match      20.4%; Score 480.5; DB 1; Length 119;
Best Local Similarity 75.0%; Pred. No. 1.1e-26;
Matches 90; Conservative 14; Mismatches 13; Indels 3; Gaps 2;

QY 142 EVLVESGGGVQPGSRSLRSLSCSASGDFDTTYMWSVRQAPGKGLWIGIHPDSSTINY 201
DB 1 EVKLSSGGGLVQPGGSLKLSLSCSASGDFDSRYMWSVRQAPGKGLWIGIHPDSSTINY 60
QY 202 APSLKDRFTISRDNKNTLFLQMSLRPDTGYVFCASL-YFGFPWFAYWGQGTPTVTSS 260
DB 61 TPSLKDKFTISRDNKNTLYLQMSKVRSEDTALYYCARLHYGY--FDVWGAGTTVTSS 118

RESULT 11
Q9QYF0
ID Q9QYF0 PRELIMINARY; PRT; 298 AA.
AC Q9QYF0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CN 8 single chain antibody.
GN Name=CN 8 scFv;
OS synthetic construct.
```



```

Query Match      19.0%; Score 445.5; DB 1; Length 126;
Best Local Similarity 71.3%; Pred. No. 3.7e-24;
Matches 92; Conservative 9; Mismatches 15; Indels 13; Gaps 3;

QY 142 EVLVESGGGVOPGRSLRLSCSASGPDFTTYMSWVRQAPGKLEWIGHDPDSSTINY 201
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 QVQLVESGGGVOPGRSLRLSCSSGFFISYMYWVRQAPGKLEWAIWDGSDQHY 60
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 202 APSLKORFTISRDNKNTLFLQMDSLRPEDTGVYFCA-----SLYGFPPWFAYWG 251
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 ADSVKGRTISRDNKNTLFLQMDSLRPEDTGVYFCARDGSGHGFCSASCFG-P--DYWG 117
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 252 QGTPVTVSS 260
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 118 QGTPVTVSS 126

RESULT 15
HV42_MOUSE
ID HV42_MOUSE STANDARD; PRT; 117 AA.
AC P01812;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Ig heavy chain V region MOPC 173.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE OF 1-104.
RX MEDLINE=72105531; PubMed=5062012;
RA Bourgois A., Fougereau M., de Preval C.;
RT "Sequence of amino acids of the NH 2 -terminal region of a mouse-
RT clonal immunoglobulin heavy chain.";
RL Eur. J. Biochem. 24:446-455 (1972).
RN [2]
RP SEQUENCE OF 105-117.
RX MEDLINE=76091933; PubMed=812695;
RA Rocca-Serra J., Milili M., Fougereau M.;
RT "Determination of the primary structure of a mouse IgG2a
RT immunoglobulin. Amino-acid sequence of the H4 cyanogen-bromide
RT fragment.";
RL Eur. J. Biochem. 59:511-523 (1975).
RN [3]
RP SEQUENCE OF 96-117 FROM N.A.
RX MEDLINE=81223769; PubMed=6787590;
RA Gough N.M., Bernard O.;
RT "Sequences of the joining region genes for immunoglobulin heavy chains
RT and their role in generation of antibody diversity.";
Proc. Natl. Acad. Sci. U.S.A. 78:509-513 (1981).
RN [4]
RP DISULFIDE BOND.
RX PubMed=11947590;
RA Bourgois A., Fougereau M.;
RT "Partial amino acid sequence of the variable region of a mouse
RT gammaG2a immunoglobulin heavy chain. Evidence for the existence of a
RT third sub-group of variability for the heavy chain pool.";
FEBS Lett. 8:285-288 (1970).
RL
CC -1- MISCELLANEOUS: This gamma-2a chain was isolated from a myeloma
CC protein.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR PIR; A91190; G2MS73.
DR HSSP; P01810; 2PBJ.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Direct protein sequencing; Immunoglobulin V region.
FT DOMAIN 1 116 Ig-like.
FT DISULFID 22 96
FT CONFLICT 105 105 N -> D (in Ref. 2).

```



FT NON TER 117 117  
SQ SEQUENCE 117 AA; 13051 MW; 156DCCC259380F19 CRC64;  
Query Match 18.9%; Score 445; DB 1; Length 117;  
Best Local Similarity 70.2%; Pred. No. 3.7e-24;  
Matches 85; Conservative 16; Mismatches 14; Indels 6; Gaps 2;  
QY 142 EVQLVESGGGVVQPGSRSLRLSCASAGDFTTYMSWVRQAPGKGLEWIGEIHPDSSTINY 201  
Db 1 EVKLLESGGPLVQLGGSLKLSCAASGDFDSRYWMSWVRQAPGKGLEWIGEIDPNSSTINY 60  
QY 202 APSLKDRFTIISRDNAKNTLFLQMSLRPEDTGVIYFCASLYFGFPWFA--YWGQGTPPTVS 259  
Db 61 TPSLKDKFTIISRDNAKNTLFLQMSKVRSEDTALYYCAR----SPYYAMNYWGQGTSTVTS 116  
QY 260 S 260  
Db 117 S 117

Search completed: May 26, 2005, 02:12:53  
Job time : 347 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 26, 2005, 01:18:54 ; Search time 95 Seconds  
(without alignments)  
348.100 Million cell updates/sec

Title: US-10-006-771B-2

Perfect score: 2350

Sequence: 1 MWSCILFLVATATGVHSD.....LSTATKDYDLMHQAALP 443

Scoring table: BLOSOM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/prodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/prodata/1/iaa/PCUTUS\_COMB.pep.\*
- 6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	965	41.1	352	4	US-09-203-958A-2
2	871	37.1	283	3	US-09-420-592A-6
3	871	37.1	283	4	US-09-985-442-6
4	871	37.1	283	4	US-09-983-580-6
5	870	37.0	263	3	US-09-069-821-3
6	870	37.0	263	4	US-09-956-086-3
7	870	37.0	263	4	US-09-956-087-3
8	838	35.7	332	4	US-09-135-121B-7
9	824	35.1	304	3	US-08-862-124-14
10	819.5	34.9	240	1	US-08-488-113B-147
11	819.5	34.9	240	1	US-08-477-484B-147
12	819.5	34.9	240	2	US-08-646-360-147
13	819.5	34.9	240	3	US-08-939-765-147
14	819.5	34.9	240	3	US-09-136-389-147
15	819.5	34.9	240	3	US-09-610-838-147
16	819.5	34.9	240	4	US-09-711-485-147
17	808	34.4	491	4	US-10-011-125A-2
18	795	33.8	553	2	US-08-661-052-16
19	795	33.8	553	3	US-09-188-082-16
20	795	33.8	553	3	US-09-364-088-16
21	795	33.8	553	3	US-09-102-716-16
22	787	33.5	354	4	US-09-393-627B-28
23	774.5	33.0	374	4	US-09-646-028-15
24	772	32.9	456	4	US-09-495-880A-11
25	769.5	32.7	265	3	US-09-420-592A-5
26	769.5	32.7	265	4	US-09-985-442-5
27	769.5	32.7	265	4	US-09-983-580-5

28	767	32.6	599	1	US-08-442-542-18	Sequence 18, Appl
29	767	32.6	599	3	US-08-765-469-18	Sequence 18, Appl
30	766.5	32.6	245	3	US-09-069-821-5	Sequence 5, Appl
31	766.5	32.6	245	4	US-09-956-086-5	Sequence 5, Appl
32	766.5	32.6	245	4	US-09-956-087-5	Sequence 5, Appl
33	765.5	32.6	241	2	US-08-224-591-18	Sequence 18, Appl
34	765.5	32.6	241	2	US-08-926-789-18	Sequence 18, Appl
35	763.5	32.5	287	3	US-08-862-124-17	Sequence 17, Appl
36	750.5	31.9	237	2	US-08-224-591-16	Sequence 16, Appl
37	750.5	31.9	237	2	US-08-926-789-16	Sequence 16, Appl
38	748	31.8	532	2	US-08-417-495-6	Sequence 6, Appl
39	748	31.8	532	2	US-08-284-391B-6	Sequence 6, Appl
40	748	31.8	532	3	US-09-218-950-6	Sequence 6, Appl
41	748	31.8	532	4	US-08-394-388A-6	Sequence 6, Appl
42	748	31.8	532	5	PCT-US92-01785-6	Sequence 6, Appl
43	748	31.8	532	5	PCT-US95-00454-6	Sequence 6, Appl
44	740	31.5	239	2	US-08-860-174A-2	Sequence 2, Appl
45	737	31.4	244	2	US-08-392-338A-13	Sequence 13, Appl

## ALIGNMENTS

### RESULT 1

US-09-203-958A-2  
; Sequence 2, Application US/09203958A  
; Patent No. 6682928

; GENERAL INFORMATION:  
; APPLICANT: KELER, Tibor  
; APPLICANT: GOLDSTEIN, Joel

; APPLICANT: GRAZIANO, Robert  
; APPLICANT: DEO, Yashwant M.

; TITLE OF INVENTION: CELLS EXPRESSING ANTI-FC RECEPTOR  
; TITLE OF INVENTION: BINDING COMPONENTS

; FILE REFERENCE: MXI-099CPA  
; CURRENT APPLICATION NUMBER: US/09/203,958A

; CURRENT FILING DATE: 1998-12-02  
; PRIOR APPLICATION NUMBER: 60/067232

; PRIOR FILING DATE: 1997-12-02  
; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2

; LENGTH: 352  
; TYPE: PRT

; ORGANISM: Artificial Sequence  
; FEATURE:

; OTHER INFORMATION: Synthetic construct  
US-09-203-958A-2

Query Match 41.1%; Score 965; DB 4; Length 352;  
Best Local Similarity 75.3%; Pred. No. 2.3e-68;  
Matches 195; Conservative 13; Mismatches 43; Indels 8; Gaps 3;

QY	12	ATATGVHSIDILQTSPSSLSASVGDRTVTTCASQDVGTIS-----VANTQQRGKAPKL	65
DB	30	AGAAQPARSDILQTSPSSLSASVGDRTVTTCSSQSVLYSSNKNYLAWTYQKPGKAPKL	89
QY	66	LIWVTSRTRTGVPSPRFGSGSGTDFTTTSSLPEDIAITYCOQSYLYRSFGQGTKEIK	125
DB	90	LIYWASTRSGVSPRFGSGSGTDFTTTSSLPEDIAITYCHQLSSWTFGGQTKVEIK	149
QY	126	-RGSGSGSGSGSGSGSGSEVQLVESGGGVVQPGKSLRSLSCSASGDFDTTVMWVRQAPGK	184
DB	150	SSGSGSGSGSGSGSGSGSEVQLVESGGGVVQPGKSLRSLSCSASGDFDTTVMWVRQAPGK	209
QY	185	GLEWIEIHPPDSTINYAPSLKDRFTISRDNAKNTLFLQWDSLRPDTGYVFCASLYFGP	244
DB	210	GLEWVATISDGGSYTYTPDSVKGRTISRDNSKNTLFLQWDSLRPDTGYVFCARGYIRY	269
QY	245	PN-FAYWGQGTPTVYSSAK	262
DB	270	EGAMDYWGQGTPTVYSSSPR	288

```
RESULT 2
US-09-420-592A-6
; Sequence 6, Application US/09420592A
; Patent No. 6333396
; GENERAL INFORMATION:
; APPLICANT: Filpula, David R.
; APPLICANT: Wang, Maoliang
; APPLICANT: Whitlow, Marc D.
; TITLE OF INVENTION: No. 6333396el Method for Targeted Delivery of Nucleic Acids
; FILE REFERENCE: 0977.2300001
; CURRENT APPLICATION NUMBER: US/09/420.592A
; CURRENT FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/104,949
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Kabat
; NAME/KEY: UNSURE
; LOCATION: (232)
; OTHER INFORMATION: May be any amino acid.
; NAME/KEY: UNSURE
; LOCATION: (234)
; OTHER INFORMATION: May be any amino acid.
; NAME/KEY: UNSURE
; LOCATION: (239)
; OTHER INFORMATION: May be any amino acid.
; US-09-420-592A-6
Query Match 37.1%; Score 871; DB 3; Length 283;
Best Local Similarity 68.2%; Pred. No. 4.7e-61;
Matches 178; Conservative 23; Mismatches 42; Indels 18; Gaps 5;
QY 20 DIQTQSPSSLSASVGDRTTTCRASQ---DVGTSVAVYQKPGKAPKLLIYWTSTRHTG 76
Db 1 DIQTQSPSSLSASVGDRTTTCRASQSLVSIISNYLAWYQKPGKAPKLLIYAASLSG 60
QY 77 VPSRFGSGSGCTDFTFISSLOPEDATYYCOQYSLYR--SFGQGTKEIKRGSGSG-- 132
Db 61 VPSRFGSGSGCTDFTFISSLOPEDATYYCOQYNSLPFWTGGQTKVEIKGSTSGSGKP 120
QY 133 GSGSGSGSEVOLVESGGVQPGRLSLSCASGDFDTTYMWSVWROAPKGLWEIGEI 192
Db 121 GSGSGSTKGEVOLVESGGVQPGRLSLSCASGDFTFSSVAMSWVROAPKGLWEIVSI 180
QY 193 --HPDSSTINYPAPSKDRFTISRDNKNTLFLQWMSLRPDTGVYFCA-----SLY 241
Db 181 SKGTDGSGTYADSVKGRFTISRDNKNTLYLQWMSLRRAEDTAVVYCARGKXGSLSGY 240
QY 242 FGFPWFAYWGQGTPTVSSAK 262
Db 241 YYHYFDYWGQGTTLVTSSKK 261
RESULT 3
US-09-985-442-6
; Sequence 6, Application US/09985442
; Patent No. 6692942
; GENERAL INFORMATION:
; APPLICANT: Filpula, David R.
; APPLICANT: Wang, Maoliang
; APPLICANT: Whitlow, Marc D.
; TITLE OF INVENTION: No. 6692942el Method for Targeted Delivery of Nucleic Acids
; FILE REFERENCE: 0977.2300003
; CURRENT APPLICATION NUMBER: US/09/985.442
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: 09/420,592
```

```
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/104,949
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Kabat
; NAME/KEY: UNSURE
; LOCATION: (232)
; OTHER INFORMATION: May be any amino acid.
; NAME/KEY: UNSURE
; LOCATION: (234)
; OTHER INFORMATION: May be any amino acid.
; NAME/KEY: UNSURE
; LOCATION: (239)
; OTHER INFORMATION: May be any amino acid.
; US-09-985-442-6
Query Match 37.1%; Score 871; DB 4; Length 283;
Best Local Similarity 68.2%; Pred. No. 4.7e-61;
Matches 178; Conservative 23; Mismatches 42; Indels 18; Gaps 5;
QY 20 DIQTQSPSSLSASVGDRTTTCRASQ---DVGTSVAVYQKPGKAPKLLIYWTSTRHTG 76
Db 1 DIQTQSPSSLSASVGDRTTTCRASQSLVSIISNYLAWYQKPGKAPKLLIYAASLSG 60
QY 77 VPSRFGSGSGCTDFTFISSLOPEDATYYCOQYSLYR--SFGQGTKEIKRGSGSG-- 132
Db 61 VPSRFGSGSGCTDFTFISSLOPEDATYYCOQYNSLPFWTGGQTKVEIKGSTSGSGKP 120
QY 133 GSGSGSGSEVOLVESGGVQPGRLSLSCASGDFDTTYMWSVWROAPKGLWEIGEI 192
Db 121 GSGSGSTKGEVOLVESGGVQPGRLSLSCASGDFTFSSVAMSWVROAPKGLWEIVSI 180
QY 193 --HPDSSTINYPAPSKDRFTISRDNKNTLFLQWMSLRPDTGVYFCA-----SLY 241
Db 181 SKGTDGSGTYADSVKGRFTISRDNKNTLYLQWMSLRRAEDTAVVYCARGKXGSLSGY 240
QY 242 FGFPWFAYWGQGTPTVSSAK 262
Db 241 YYHYFDYWGQGTTLVTSSKK 261
RESULT 4
US-09-983-580-6
; Sequence 6, Application US/09983580
; Patent No. 6764853
; GENERAL INFORMATION:
; APPLICANT: Filpula, David R.
; APPLICANT: Wang, Maoliang
; APPLICANT: Whitlow, Marc D.
; TITLE OF INVENTION: No. 6764853el Method for Targeted Delivery of Nucleic Acids
; FILE REFERENCE: 0977.2300002
; CURRENT APPLICATION NUMBER: US/09/983.580
; CURRENT FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 09/420,592
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/104,949
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Kabat
; OTHER INFORMATION: Consensus
```

```

; NAME/KEY: UNSURE
; LOCATION: (232)
; OTHER INFORMATION: May be any amino acid.
; NAME/KEY: UNSURE
; LOCATION: (234)
; OTHER INFORMATION: May be any amino acid.
; NAME/KEY: UNSURE
; LOCATION: (239)
; OTHER INFORMATION: May be any amino acid.
US-09-983-580-6

Query Match      37.18; Score 871; DB 4; Length 283;
Best Local Similarity 68.28; Pred. No. 4.7e-61;
Matches 178; Conservative 23; Mismatches 42; Indels 18; Gaps 5;

QY 20 DIQTQSPSSLSASVGRVITTCASQ---DVGTSVAVYQKPKAPKLLIYWTSTRHTG 76
Db 1 DIQTQSPSSLSASVGRVITTCASQSLVISNYLAWYQKPKAPKLLIYAASLSLG 60

QY 77 VPSRFGSGSGTDTFTTISLQPEDIATYCCQYSLYR--SPGGTKVEIKRGGSGG-- 132
Db 61 VPSRFGSGSGTDTFTTISLQPEDFATYCCQYNSLPWTFGTKVEIKGSTSGSGKP 120

QY 133 GSGSGGSGEVQVSGGVQPGSRSLRSCSAGDFFTTYMWSVWRQAPKGLWIGEI 192
Db 121 GSGEGSTKGEVQVSGGLVQPGGSLRLSCAASGFTFSSYAMSWVRQAPKGLWYSVI 180

QY 193 --HPDSSTINYAPSLKDRFTISRDNKNTLFLQMSLRPDTGVYFCA-----SLY 241
Db 181 SKTDCGSGTYADSVKGRFTISRDNKNTLYLQNSLRABDTAVYYCARGXKXSLSGXY 240

QY 242 FGPFWFAYWQGTPTVTSSAK 262
Db 241 YYHYFDYWGQGLTVTVSSKK 261

RESULT 5
US-09-069-821-3
; Sequence 3, Application US/09069821
; Patent No. 632322
; GENERAL INFORMATION:
; APPLICANT: FILPULA, DAVID
; APPLICANT: WANG, MAOLIANG
; APPLICANT: SHORR, ROBERT
; APPLICANT: WHITLOW, MARC
; APPLICANT: LEE, LIHSYNG S.
; TITLE OF INVENTION: SINGLE-CHAIN ANTIGEN-BINDING PROTEINS
; CAPABLE OF GLYCOSYLATION, PRODUCTION AND USES THEREOF
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/069,821
; FILING DATE: 30-APR-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/067,341
; FILING DATE: 02-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/063,074
; FILING DATE: 27-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/050,472

```

```

; FILING DATE: 23-JUN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/044,449
; FILING DATE: 30-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: KIM, JUDITH U.
; REGISTRATION NUMBER: 40,679
; REFERENCE/DOCKET NUMBER: 0977.2280003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)371-2600
; TELEFAX: (202)371-2540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-09-069-821-3

Query Match      37.0%; Score 870; DB 3; Length 263;
Best Local Similarity 68.2%; Pred. No. 5.2e-61;
Matches 178; Conservative 23; Mismatches 42; Indels 18; Gaps 5;

QY 20 DIQTQSPSSLSASVGRVITTCASQ---DVGTSVAVYQKPKAPKLLIYWTSTRHTG 76
Db 1 DIQTQSPSSLSASVGRVITTCASQSLVISNYLAWYQKPKAPKLLIYAASLSLG 60

QY 77 VPSRFGSGSGTDTFTTISLQPEDIATYCCQYSLYR--SFGGTKVEIKRGGSGG-- 132
Db 61 VPSRFGSGSGTDTFTTISLQPEDFATYCCQYNSLPWTFGTKVEIKGSTSGSGKP 120

QY 133 GSGSGGSGEVQVSGGVQPGSRSLRSCSAGDFFTTYMWSVWRQAPKGLWIGEI 192
Db 121 GSGEGSTKGEVQVSGGLVQPGGSLRLSCAASGFTFSSYAMSWVRQAPKGLWYSVI 180

QY 193 --HPDSSTINYAPSLKDRFTISRDNKNTLFLQMSLRPDTGVYFCA-----SLY 241
Db 181 SKTDCGSGTYADSVKGRFTISRDNKNTLYLQNSLRABDTAVYYCARGXKXSLSGXY 240

QY 242 FGPFWFAYWQGTPTVTSSAK 262
Db 241 YYHYFDYWGQGLTVTVSSNK 261

RESULT 6
US-09-956-086-3
; Sequence 3, Application US/09956086
; Patent No. 6743896
; GENERAL INFORMATION:
; APPLICANT: FILPULA, DAVID
; APPLICANT: WANG, MAOLIANG
; APPLICANT: SHORR, ROBERT
; APPLICANT: WHITLOW, MARC
; APPLICANT: LEE, LIHSYNG S.
; TITLE OF INVENTION: SINGLE-CHAIN ANTIGEN-BINDING PROTEINS
; CAPABLE OF GLYCOSYLATION, PRODUCTION AND USES THEREOF
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/956,086
; FILING DATE: 20-Sep-2001

```

## CLASSIFICATION: &lt;Unknown&gt;

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/069,821  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US 60/063,074  
FILING DATE: 27-OCT-1997  
APPLICATION NUMBER: US 60/050,472  
FILING DATE: 23-JUN-1997  
APPLICATION NUMBER: US 60/044,449  
FILING DATE: 30-APR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: KIM, JUDITH U.  
REGISTRATION NUMBER: 40,679  
REFERENCE/DOCKET NUMBER: 0977.2280003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)371-2600  
TELEFAX: (202)371-2540  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 263 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: No. 6743896 Relevant  
MOLECULE TYPE: Peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-956-086-3

Query Match 37.0%; Score 870; DB 4; Length 263;  
Best Local Similarity 68.2%; Pred. No. 5.2e-61;  
Matches 178; Conservative 23; Mismatches 42; Indels 18; Gaps 5;

QY 20 DIQQTSPSSLSASVGRVITTCASQ---DVGTSAVYQKPKGKAPKLLIYWTSTRHTG 76  
DB 1 DIQQTSPSSLSASVGRVITTCASQSLVSIINLAWYQKPKGKAPKLLIYAASSLSG 60

QY 77 VPSRFGSGSGTDFTTISSLOPEDFATYYCQYSLYR--SFGQGTKEIKRGSGSG-- 132  
DB 61 VPSRFGSGSGTDFTTISSLOPEDFATYYCQYNSLPFWTGGQTKVEIKGTSGSGKP 120

QY 133 GSGGSGSGSEVQLVESGGVQVPGKSLRLSCASGFDFTTVMWVQRQAPKGLWIGEI 192  
DB 121 GSGEGSTKGEVQLVESGGGLVQPGGSLRLSCAASGFTFSSYAMSVQRQAPKGLWVSVI 180

QY 193 --HPDSSTINAPSLKDRFTISRDNKNTLFLOMDSLRLPDTGVYFCA-----SLY 241  
DB 181 SKGTDGGSTYYADSVKGRFTISRDNKNTLYLQWNSLRAEDTAVYYCARGRXGSLSGY 240

QY 242 FGFPWFAYWGQGTPTVVSNAK 262  
DB 241 YYHYFDYWGQGTTLTVSSNK 261

## RESULT 7

US-09-956-087-3  
Sequence 3, Application US/09956087  
Patent No. 6743908

## GENERAL INFORMATION:

APPLICANT: FILIPULA, DAVID

WANG, MAOLIANG

SHORR, ROBERT

WHITLOW, MARC

LEE, LIHSYNG S.

TITLE OF INVENTION: SINGLE-CHAIN ANTIGEN-BINDING PROTEINS

CAPABLE OF GLYCOSYLATION, PRODUCTION AND USES THEREOF

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESS: STERNE, KESSLER, GOLDSTEIN &amp; FOX P.L.L.C.

STREET: 1100 NEW YORK AVE., NW, SUITE 600

CITY: WASHINGTON

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/956,087  
FILING DATE: 20-Sep-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/069,821  
FILING DATE: 1998-04-30  
APPLICATION NUMBER: US 60/063,074  
FILING DATE: 27-OCT-1997  
APPLICATION NUMBER: US 60/050,472  
FILING DATE: 23-JUN-1997  
APPLICATION NUMBER: US 60/044,449  
FILING DATE: 30-APR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: KIM, JUDITH U.  
REGISTRATION NUMBER: 40,679  
REFERENCE/DOCKET NUMBER: 0977.2280003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)371-2600  
TELEFAX: (202)371-2540  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 263 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-956-087-3

Query Match 37.0%; Score 870; DB 4; Length 263;

Best Local Similarity 68.2%; Pred. No. 5.2e-61;

Matches 178; Conservative 23; Mismatches 42; Indels 18; Gaps 5;

QY 20 DIQQTSPSSLSASVGRVITTCASQ---DVGTSAVYQKPKGKAPKLLIYWTSTRHTG 76  
DB 1 DIQQTSPSSLSASVGRVITTCASQSLVSIINLAWYQKPKGKAPKLLIYAASSLSG 60

QY 77 VPSRFGSGSGTDFTTISSLOPEDFATYYCQYSLYR--SFGQGTKEIKRGSGSG-- 132  
DB 61 VPSRFGSGSGTDFTTISSLOPEDFATYYCQYNSLPFWTGGQTKVEIKGTSGSGKP 120

QY 133 GSGGSGSGSEVQLVESGGVQVPGKSLRLSCASGFDFTTVMWVQRQAPKGLWIGEI 192  
DB 121 GSGEGSTKGEVQLVESGGGLVQPGGSLRLSCAASGFTFSSYAMSVQRQAPKGLWVSVI 180

QY 193 --HPDSSTINAPSLKDRFTISRDNKNTLFLOMDSLRLPDTGVYFCA-----SLY 241  
DB 181 SKGTDGGSTYYADSVKGRFTISRDNKNTLYLQWNSLRAEDTAVYYCARGRXGSLSGY 240

QY 242 FGFPWFAYWGQGTPTVVSNAK 262  
DB 241 YYHYFDYWGQGTTLTVSSNK 261

## RESULT 8

US-09-135-121B-7

Sequence 7, Application US/09135121B

Patent No. 6534051

GENERAL INFORMATION:

APPLICANT: Dornburg, Ralph C.

TITLE OF INVENTION: CELL TYPE SPECIFIC GENE TRANSFER USING RETROVIRAL VECTORS

TITLE OF INVENTION: CONTAINING ANTIBODY-ENVELOPE FUSION PROTEINS AND WILD-TYPE

FILE REFERENCE: 97,216-L

CURRENT APPLICATION NUMBER: US/09/135,121B

CURRENT FILING DATE: 1998-08-17

PRIOR APPLICATION NUMBER: US 08/933,616

PRIOR FILING DATE: 1997-08-28

PRIOR APPLICATION NUMBER: US 08/205,980

```
; PRIOR FILING DATE: 1994-03-04
; PRIOR APPLICATION NUMBER: US 07/979,619
; PRIOR FILING DATE: 1992-11-20
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Microsoft Word 97
; SEQ ID NO 7
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Anti-Her2neu single chain antibody
US-09-135-121B-7

Query Match          35.7%; Score 838; DB 4; Length 332;
Best Local Similarity 48.5%; Pred. No. 2.4e-58;
Matches 180; Conservative 53; Mismatches 76; Indels 62; Gaps 10;

QY 6 IILFLVATATGVHS--DIQLTQSPSSLSASVGDRTVITCKASQDVGTSAVWYQKPKAP 63
DB 7 IFSLFLISASVIASRGDIVATQPKFASTSVGDRISVTCKAS-DVGNVAWYQKPKQSP 65

QY 64 KLIYWTSTRTHTGVPSRFGSGSGTDFTTISLQPEDIAITYYCOQVSLYR-SFGQGTKV 122
DB 66 KPIIYSASYLYNGVPRFTGSGSGTDFSLISNVQSDDLAEYFCQYNTYPTFFGGTKL 125

QY 123 EIKRGSGSGSGSGSGSGVQVLESQVGVVQPGSRSLRSCSASGDFFTTYWMSVWRQAP 182
DB 126 EIKRGSGSGKSSEG--KGEVQLEESGGGLVQPKGSLKLSCAASGFTFTNTYAANVWRQAP 183

QY 183 GKLEWIGETHPDSTINYA-----PSLKDRFTTISRDNAKNTLFLQWDSLRLPDTGYFCA 238
DB 184 GKLEWIVIRIRKSN--NYATYVYDYSKDRFTTISRDNSQALYLOANLKTEDTAAYYCV 241

QY 239 SLVFGEP--WFAYWGQGTPTVSSAKPTTTPAPRPPTPTPTIASQPLSLRPEARPAAGG 296
DB 242 TSIDYDKVLFAWGQGTPTVSSA-----266

QY 297 AVHTRGLDFALDKPLCYLLGILFYGVILTALFLRVKFS-RSAEPPAYQOQGNLYNEL 355
DB 267 -----DQQLCYLLDAILFLYGLVLTLYCLRLKIQVRKAITSYEKSDG-VYTGL 314

QY 356 NLGRREYDVL 366
DB 315 STRNQETYTEL 325

RESULT 9
US-08-862-124-14
; Sequence 14, Application US/08862124
; Patent No. 6207153
; GENERAL INFORMATION:
; APPLICANT: Dan, Michael D.
; APPLICANT: Maiti, Pradip K.
; APPLICANT: Kaplan, Howard A.
; TITLE OF INVENTION: ANTIGEN BINDING FRAGMENTS H11, THAT
; TITLE OF INVENTION: SPECIFICALLY DETECT CANCER CELLS, NUCLEOTIDES ENCODING THE
; TITLE OF INVENTION: FRAGMENTS, AND USE THEREOF FOR THE PROPHYLAXIS AND
; TITLE OF INVENTION: DETECTION OF CANCERS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster LLP
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/862,124
```

```
; FILING DATE: 22-MAY-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Lehnhardt, Susan K.
; REGISTRATION NUMBER: 33,943
; REFERENCE/DOCKET NUMBER: 31608-20001.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 813-5600
; TELEFAX: (650) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 304 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-862-124-14

Query Match          35.1%; Score 824; DB 3; Length 304;
Best Local Similarity 61.9%; Pred. No. 2.7e-57;
Matches 172; Conservative 29; Mismatches 51; Indels 26; Gaps 6;

QY 6 IILFLVATATGVHS--DIQLTQSPSSLSASVGDRTVITCKASQDVGTSAVWYQKPKAP 64
DB 10 IAVLAGPATVAQADIVLTQSPGTLSPGERATLSCRAQSQSVSSYLAWYQKPKQAPR 69

QY 65 LLIYWTSTRTHTGVPSRFGSGSGTDFTTISLQPEDIAITYYCOQV-----SLYRSFGQG 119
DB 70 LLIYGASTRATGMPDRFSGSGSGTDFTLTISRLEPEDFAVYVCOQYSGSSPQTPTITFGGG 129

QY 120 TKVEIKR-----GGSGSGSGSGSGSEVOLVESGGVQVQPGSRSLRSCSASGDFDT 171
DB 130 TKVEIKRTVAAPSVSQSGSGSGSGGQVQVLESQVGVVQPGSRSLRSCSASGDFPR 189

QY 172 TYWMSVWRQAPGKLEWIGETHPDSTINYAPSLKDRFTTISRDNAKNTLFLQWDSLRLP 231
DB 190 SPAMHWVRQALGKLEWAVISYDGTSTKYADSVKGRFTISRDTSKNTVTLKNSLTED 249

QY 232 TGVYFCA---SL-----YFGFPEFAYWGQGTPTVSS 260
DB 250 TAVYVCARDQSLGDDHYHYGLD---VMGKGTITVTVSS 284

RESULT 10
US-08-488-113B-147
; Sequence 147, Application US/08488113B
; Patent No. 5744580
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Studnika, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 169
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,113B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/425,336
; FILING DATE: 18-APR-1995
```



```
/
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/064,691
/ FILING DATE: 12-MAY-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/988,430
/ FILING DATE: 09-DEC-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/901,707
/ FILING DATE: 19-JUN-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/787,567
/ FILING DATE: 04-NOV-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: McNicholas, Janet M.
/ REGISTRATION NUMBER: 32,918
/ REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 312/707-8889
/ TELEFAX: 312/707-9155
/ TELEX: 650 388-1248
/ INFORMATION FOR SEQ ID NO: 147:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 240 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-488-113B-147

Query Match 34.9%; Score 819.5; DB 1; Length 240;
Best Local Similarity 65.8%; Pred. No. 4.5e-57;
Matches 160; Conservative 31; Mismatches 47; Indels 5; Gaps 4;

QY 20 DIQTQSPSSLSASVGRVTTICKASQDVGTSAVYQOKPGKAPKLLIYWTSTRHTGVPS 79
Db 1 DIQMTQSPSSLSASVGRVTTICKASQDVGTSAVYQOKPGKAPKLLIYWTSTRHTGVPS 60

QY 80 RFGSGSGTDYTLTISSLYQYEDFGIYCCQYDESPWTFGGTKLEMK-GGGSGGGSGG 138
Db 61 RFGSGSGTDYTLTISSLYQYEDFGIYCCQYDESPWTFGGTKLEMK-GGGSGGGSGG 119

QY 139 SGSEVQLVESGGGVVQPGSRSLRSCASGDFPTTVMNSVWVROAPKGLIEWGIEIHPDSST 198
Db 139 SGSEVQLVESGGGVVQPGSRSLRSCASGDFPTTVMNSVWVROAPKGLIEWGIEIHPDSST 198

QY 120 GGSEIQLVQSGGGLVKPGGSRVISAASGYFTTNYGMNWNVROAPKGLIEWGINTHTGE 179
Db 120 GGSEIQLVQSGGGLVKPGGSRVISAASGYFTTNYGMNWNVROAPKGLIEWGINTHTGE 179

QY 199 INYAPSLKDRFTISRDNKNTLFLQMSLRPEDTGVCASLYFGFPW-FAYWGGGTPTV 257
Db 180 PTYADSFKGRFTSLDSSKNTAYLIQINSRAEDTAVYFCTRR--GYDWFYDVWGGGTPTV 237

QY 258 VSS 260
Db 238 VSS 240

RESULT 11
US-08-477-484B-147
/ Sequence 147, Application US/08477484B
/ Patent No. 5756659
/ GENERAL INFORMATION:
/ APPLICANT: Better, Marc D.
/ APPLICANT: Carroll, Stephen F.
/ APPLICANT: Studnika, Gary M.
/ TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
/ NUMBER OF SEQUENCES: 169
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: McAndrews, Held & Malloy, Ltd.
/ STREET: 500 West Madison Street, 34th floor
/ CITY: Chicago
/ STATE: Illinois
/ COUNTRY: USA
/ ZIP: 60661
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk

/
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/477,484B
/ FILING DATE: 07-JUN-1995
/ CLASSIFICATION: 530
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/425,336
/ FILING DATE: 18-APR-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/064,691
/ FILING DATE: 12-MAY-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/988,430
/ FILING DATE: 09-DEC-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/901,707
/ FILING DATE: 19-JUN-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/787,567
/ FILING DATE: 04-NOV-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: McNicholas, Janet M.
/ REGISTRATION NUMBER: 32,918
/ REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 312/707-8889
/ TELEFAX: 312/707-9155
/ TELEX: 650 388-1248
/ INFORMATION FOR SEQ ID NO: 147:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 240 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-477-484B-147

Query Match 34.9%; Score 819.5; DB 1; Length 240;
Best Local Similarity 65.8%; Pred. No. 4.5e-57;
Matches 160; Conservative 31; Mismatches 47; Indels 5; Gaps 4;

QY 20 DIQTQSPSSLSASVGRVTTICKASQDVGTSAVYQOKPGKAPKLLIYWTSTRHTGVPS 79
Db 1 DIQMTQSPSSLSASVGRVTTICKASQDVGTSAVYQOKPGKAPKLLIYWTSTRHTGVPS 60

QY 80 RFGSGSGTDYTLTISSLYQYEDFGIYCCQYDESPWTFGGTKLEMK-GGGSGGGSGG 138
Db 61 RFGSGSGTDYTLTISSLYQYEDFGIYCCQYDESPWTFGGTKLEMK-GGGSGGGSGG 119

QY 139 SGSEVQLVESGGGVVQPGSRSLRSCASGDFPTTVMNSVWVROAPKGLIEWGIEIHPDSST 198
Db 139 SGSEVQLVESGGGVVQPGSRSLRSCASGDFPTTVMNSVWVROAPKGLIEWGIEIHPDSST 198

QY 120 GGSEIQLVQSGGGLVKPGGSRVISAASGYFTTNYGMNWNVROAPKGLIEWGINTHTGE 179
Db 120 GGSEIQLVQSGGGLVKPGGSRVISAASGYFTTNYGMNWNVROAPKGLIEWGINTHTGE 179

QY 199 INYAPSLKDRFTISRDNKNTLFLQMSLRPEDTGVCASLYFGFPW-FAYWGGGTPTV 257
Db 180 PTYADSFKGRFTSLDSSKNTAYLIQINSRAEDTAVYFCTRR--GYDWFYDVWGGGTPTV 237

QY 258 VSS 260
Db 238 VSS 240

RESULT 12
US-08-646-360-147
/ Sequence 147, Application US/08646360
/ Patent No. 5837491
/ GENERAL INFORMATION:
/ APPLICANT: Better, Marc D.
/ APPLICANT: Carroll, Stephen F.
/ APPLICANT: Studnika, Gary M.
/ TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
/ TITLE OF INVENTION: Proteins
```

NUMBER OF SEQUENCES: 173  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,360  
FILING DATE: 13-MAY-1996  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/05348  
FILING DATE: 12-MAY-1994  
APPLICATION DATA:  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 200-70.P4  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 147:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 240 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-646-360-147

Query Match 34.9%; Score 819.5; DB 2; Length 240;  
Best Local Similarity 65.8%; Pred. No. 4.5e-57;  
Matches 160; Conservative 31; Mismatches 47; Indels 5; Gaps 4;  
Qy 20 DIQLTQSPSLASVGRVITTCASQDVGTSVAVYQOKPKAPKLLIYWTSTRHTGVPS 79  
Db 1 DIQMTQSPSLASVGRVITTCRASQDINSYLSWFQOKPKAPKLLIYRANRLESGVPS 60  
Qy 80 RFGSGSGTDFTTISLQPEDIATYYCQYSLYR-SFGQGTKEIKRGGSGSGSGG 138  
Db 61 RFGSGSGTDYTLTISLQYEDFGIYYCQYDPSPTFGGTGKLEMK-GGGSGSGSGG 119  
Qy 139 SGSEVQLVSGGGVQPGRLSLRSCSASGFDTFTYMWVSWVRQAPGKGLWIGIHPDSST 198  
Db 120 GGSEIQLVQSGGLVPGGVSRIASCAAGTFTNYGMNVVRQAPGKGLWGWINTHTGE 179  
Qy 199 INYAPSLKDRFTTISRDNAKTLFLQMDSLRPEDTGVYFCASLFGFPW-FAYWGQGTPT 257  
Db 180 PTYADSPKGRFTSLDDSKNTAYLQINSRAEDTAVYFCTRR--GYDWFVDMGQGTVT 237  
Qy 258 VSS 260  
Db 238 VSS 240

RESULT 13  
US-08-839-765-147  
Sequence 147, Application US/08839765  
Patent No. 6146631  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
TITLE OF INVENTION: Proteins  
NUMBER OF SEQUENCES: 169  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/839,765  
FILING DATE: 15-APR-1997  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/425,336  
FILING DATE: 18-APR-1995  
APPLICATION DATA:  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 11022US09/200-70.P3.C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 147:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 240 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-839-765-147

Query Match 34.9%; Score 819.5; DB 3; Length 240;  
Best Local Similarity 65.8%; Pred. No. 4.5e-57;  
Matches 160; Conservative 31; Mismatches 47; Indels 5; Gaps 4;  
Qy 20 DIQLTQSPSLASVGRVITTCASQDVGTSVAVYQOKPKAPKLLIYWTSTRHTGVPS 79  
Db 1 DIQMTQSPSLASVGRVITTCRASQDINSYLSWFQOKPKAPKLLIYRANRLESGVPS 60  
Qy 80 RFGSGSGTDFTTISLQPEDIATYYCQYSLYR-SFGQGTKEIKRGGSGSGSGG 138  
Db 61 RFGSGSGTDYTLTISLQYEDFGIYYCQYDPSPTFGGTGKLEMK-GGGSGSGSGG 119  
Qy 139 SGSEVQLVSGGGVQPGRLSLRSCSASGFDTFTYMWVSWVRQAPGKGLWIGIHPDSST 198  
Db 120 GGSEIQLVQSGGLVPGGVSRIASCAAGTFTNYGMNVVRQAPGKGLWGWINTHTGE 179



TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 147:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 240 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-610-838-147

Query Match 34.9%; Score 819.5; DB 3; Length 240;  
Best Local Similarity 65.8%; Pred. No. 4.5e-57;  
Matches 160; Conservative 31; Mismatches 47; Indels 5; Gaps 4;  
Qy 20 DIQLTQSPSSLASVGRDVTITCKASQDVGTSTVAWYQQKPKAPKLLIYWTSTRHTGVPS 79  
Db 1 DIQMTQSPSSLASVGRDVTITCEASQDINSYLSWFQKPKAPKTLIYRANRLESGVPS 60  
Qy 80 RFGSGSGTDFTFTISSLPEDIAIYYCQOYSLYR-SFGQGTKEIKRGGSGSGSGGG 138  
Db 61 RFGSGSGTDYTLTIISSLQYEDFGIYYCQYDESPWTFGGGTKEMK-GGGSGSGSGGG 119  
Qy 139 SGSEVQLVESGGGVQPGSRSLRSCSASGPDFTTYMNSWYRQAPGKLEWIGEIHPDSST 198  
Db 120 GGSEIQLVQSGGGLVPGGSGVRISCAASGYFTNYGMNWRQAPGKLEWGWINTHTGE 179  
Qy 199 INYAPSLKDRFTISRDNKNTLFLQMDSLRPEDTGVYFCASLYFGPW-FAYWGQGTPTV 257  
Db 180 PTYADSFKGRFTSLDSDSKNTAYLIQINSRAEDTAVYFCTRR--GYDWYFDVWGQGTTVT 237  
Qy 258 VSS 260  
Db 238 VSS 240

Search completed: May 26, 2005, 02:16:51  
Job time : 96 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 26, 2005, 02:12:59 ; Search time 337 Seconds  
(without alignments)  
439.724 Million cell updates/sec

Title: US-10-006-771b-2

Perfect score: 2350

Sequence: 1 MGWSCIILFLVATATGVHSD.....LSTATKDYDALHMQALPFR 443

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1434725 seqs, 334507595 residues

Total number of hits satisfying chosen parameters: 1434725

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2350	100.0	443	13	US-10-006-773-2
2	2350	100.0	443	13	US-10-006-771a-2
3	1228	52.3	634	16	US-10-416-011-2
4	1205.5	51.3	364	16	US-10-829-388-11
5	1205.5	51.3	370	16	US-10-829-388-1
6	1202.5	51.2	358	16	US-10-829-388-12
7	1202.5	51.2	363	16	US-10-829-388-2
8	1094	46.6	444	8	US-08-812-393A-2
9	1092	46.5	449	10	US-09-774-681-2
10	965	41.1	352	10	US-09-203-958A-2
11	910	38.7	631	15	US-10-120-198B-2
12	871	37.1	283	9	US-09-983-580-6
13	871	37.1	283	9	US-09-985-442-6

14	870	37.0	263	9	US-09-956-086-3	Sequence 3, Appli
15	870	37.0	263	9	US-09-956-087-3	Sequence 3, Appli
16	865.5	36.8	247	15	US-10-423-847-13	Sequence 13, Appl
17	865.5	36.8	247	17	US-10-831-063-13	Sequence 13, Appl
18	865.5	36.8	252	15	US-10-423-847-10	Sequence 10, Appl
19	865.5	36.8	252	17	US-10-831-063-10	Sequence 10, Appl
20	865.5	36.8	253	15	US-10-423-847-11	Sequence 11, Appl
21	865.5	36.8	253	17	US-10-831-063-11	Sequence 11, Appl
22	865.5	36.8	254	15	US-10-423-847-17	Sequence 17, Appl
23	865.5	36.8	254	17	US-10-831-063-17	Sequence 17, Appl
24	862.5	36.7	252	15	US-10-423-847-14	Sequence 14, Appl
25	862.5	36.7	252	17	US-10-831-063-14	Sequence 14, Appl
26	846	36.0	239	15	US-10-423-847-18	Sequence 18, Appl
27	846	36.0	239	17	US-10-831-063-18	Sequence 18, Appl
28	842.5	35.9	252	9	US-09-971-543-1	Sequence 1, Appli
29	838	35.7	332	14	US-10-211-488-7	Sequence 7, Appli
30	824	35.1	304	10	US-09-782-397-14	Sequence 14, Appl
31	824	35.1	304	15	US-10-651-453-14	Sequence 14, Appl
32	821	34.9	245	17	US-10-864-818-1	Sequence 1, Appli
33	819.5	34.9	240	14	US-10-127-890-147	Sequence 147, App
34	819.5	34.9	240	17	US-10-717-243-147	Sequence 147, App
35	808	34.4	491	13	US-10-011-125-2	Sequence 2, Appli
36	799.5	34.0	550	14	US-10-207-655-270	Sequence 270, App
37	780	33.2	266	14	US-10-207-655-260	Sequence 260, App
38	775	33.0	352	16	US-10-333-235A-54	Sequence 54, Appl
39	775	33.0	468	16	US-10-333-235A-56	Sequence 56, Appl
40	775	33.0	507	14	US-10-074-596-11	Sequence 11, Appl
41	775	33.0	565	16	US-10-333-235A-57	Sequence 57, Appl
42	774.5	33.0	374	14	US-10-335-394-15	Sequence 15, Appl
43	774	32.9	502	16	US-10-679-620-88	Sequence 88, Appl
44	772	32.9	456	15	US-10-634-862-11	Sequence 11, Appl
45	770.5	32.8	4852	15	US-10-412-406-33	Sequence 33, Appl

ALIGNMENTS

RESULT 1

US-10-006-773-2  
; Sequence 2, Application US/10006773  
; Publication No. US20020132983A1  
; GENERAL INFORMATION:  
; APPLICANT: Jungsans, Richard P.  
; TITLE OF INVENTION: Antibodies as Chimeric Effector Cell Receptors Against Tumor Anti  
; FILE REFERENCE: 003  
; CURRENT APPLICATION NUMBER: US/10/006,773  
; CURRENT FILING DATE: 2001-12-10  
; PRIOR APPLICATION NUMBER: 60/250,089  
; PRIOR FILING DATE: 2000-11-30  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 2  
; LENGTH: 443  
; TYPE: PRT  
; ORGANISM: Homo sapiens and Mus sp.  
US-10-006-773-2

Query Match 100.0%; Score 2350; DB 13; Length 443;  
Best Local Similarity 100.0%; Pred. No. 9.6e-150;  
Matches 443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGWSCIILFLVATATGVHSDIQLTQSPSSLSASVGDVITTCASQDVGTSVAWYQKPG	60
Db	1	MGWSCIILFLVATATGVHSDIQLTQSPSSLSASVGDVITTCASQDVGTSVAWYQKPG	60
Qy	61	KAPKLIYWTSTRTHTGVPRFSGSGTDFTTISLQPEDIATYCCQYSLYRSFGGT	120
Db	61	KAPKLIYWTSTRTHTGVPRFSGSGTDFTTISLQPEDIATYCCQYSLYRSFGGT	120
Qy	121	KVEIKRGGSGSGSGSGSGSEVOLYESGGVGVQPGSLRLSCSASGDFDTTYMMSVQR	180
Db	121	KVEIKRGGSGSGSGSGSGSEVOLYESGGVGVQPGSLRLSCSASGDFDTTYMMSVQR	180



```

Qy 181 APKGLEWGEIHPDSSINIVAPSLKDRFTISRDNKNTLFLQMSLRPBDTGVYFCASL 240
Db 181 APKGLEWGEIHPDSSINIVAPSLKDRFTISRDNKNTLFLQMSLRPBDTGVYFCASL 240

Qy 241 YFGPWFAYWGGTPTVTVSSAKPTTTPAPRPTTAPTASQPLSRPAAAPAGAVHT 300
Db 241 YFGPWFAYWGGTPTVTVSSAKPTTTPAPRPTTAPTASQPLSRPAAAPAGAVHT 300

Qy 301 RGLDPALDPKCLYLDGLIFLYGVLITLALFLRVKFSRABPPAYOQGNOLYNELNLR 360
Db 301 RGLDPALDPKCLYLDGLIFLYGVLITLALFLRVKFSRABPPAYOQGNOLYNELNLR 360

Qy 361 BEYDVKRRGRDPEMGKPRKPNQEGLYNELQDKMAEYSEIGMKGERRRGKHGDL 420
Db 361 BEYDVKRRGRDPEMGKPRKPNQEGLYNELQDKMAEYSEIGMKGERRRGKHGDL 420

Qy 421 YQGLSTATKOTYDALHMOALPPR 443
Db 421 YQGLSTATKOTYDALHMOALPPR 443

RESULT 2
US-10-006-771A-2
; Sequence 2, Application US/10006771A
; Publication No. US20020165360A1
; GENERAL INFORMATION:
; APPLICANT: Junghans, Richard P.
; TITLE OF INVENTION: Chimeric Effector Cell Receptors Against Carcinoembryonic Antigen
; FILE REFERENCE: 002
; CURRENT APPLICATION NUMBER: US/10/006,771A
; PRIOR APPLICATION NUMBER: 2002-06-04
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-006-771A-2

```

```

Query Match 100.0%; Score 2350; DB 13; Length 443;
Best Local Similarity 100.0%; Pred. No. 9.6e-150;
Matches 443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MWSCIIILFLVATATGVHSDIQLTQSPSSLSASVGDRTITCKASQDVGTSVAMVYQKPG 60
Db 1 MWSCIIILFLVATATGVHSDIQLTQSPSSLSASVGDRTITCKASQDVGTSVAMVYQKPG 60

Qy 61 KAPKLLIYWTSTRTHTGVPSRFGSGSGTDTFTTISLQPEDIAITYCOOYSLYRSFGGOT 120
Db 61 KAPKLLIYWTSTRTHTGVPSRFGSGSGTDTFTTISLQPEDIAITYCOOYSLYRSFGGOT 120

Qy 121 KVEIKRGGSGGGSGGSEVQLVESGGVQVQGRSLRSLSCSAGDFDTTYWMSWRQ 180
Db 121 KVEIKRGGSGGGSGGSEVQLVESGGVQVQGRSLRSLSCSAGDFDTTYWMSWRQ 180

Qy 181 APKGLEWGEIHPDSSINIVAPSLKDRFTISRDNKNTLFLQMSLRPBDTGVYFCASL 240
Db 181 APKGLEWGEIHPDSSINIVAPSLKDRFTISRDNKNTLFLQMSLRPBDTGVYFCASL 240

Qy 241 YFGPWFAYWGGTPTVTVSSAKPTTTPAPRPTTAPTASQPLSRPAAAPAGAVHT 300
Db 241 YFGPWFAYWGGTPTVTVSSAKPTTTPAPRPTTAPTASQPLSRPAAAPAGAVHT 300

Qy 301 RGLDPALDPKCLYLDGLIFLYGVLITLALFLRVKFSRABPPAYOQGNOLYNELNLR 360
Db 301 RGLDPALDPKCLYLDGLIFLYGVLITLALFLRVKFSRABPPAYOQGNOLYNELNLR 360

Qy 361 BEYDVKRRGRDPEMGKPRKPNQEGLYNELQDKMAEYSEIGMKGERRRGKHGDL 420
Db 361 BEYDVKRRGRDPEMGKPRKPNQEGLYNELQDKMAEYSEIGMKGERRRGKHGDL 420

```

```

Qy 421 YQGLSTATKOTYDALHMOALPPR 443
Db 421 YQGLSTATKOTYDALHMOALPPR 443

RESULT 3
US-10-416-011-2
; Sequence 2, Application US/10416011
; Publication No. US20040126363A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Michael
; APPLICANT: Forman, Stephen
; APPLICANT: Raubitschek, Andrew
; TITLE OF INVENTION: CD19-specific redirected immune cells
; FILE REFERENCE: 1954-338
; CURRENT APPLICATION NUMBER: US/10/416,011
; CURRENT FILING DATE: 2003-05-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 634
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: CD19R: zeta chimeric receptor
US-10-416-011-2

Query Match 52.3%; Score 1228; DB 16; Length 634;
Best Local Similarity 45.5%; Pred. No. 3.1e-74;
Matches 279; Conservative 53; Mismatches 91; Indels 190; Gaps 15;

Qy 20 DIQLTQSPSSLSASVGDRTITCKASQDVGTSVAMVYQKPGKAPKLLIYWTSTRTHTGVPS 79
Db 23 DIQMTQTSSLSASIGDRVTIISCRASQDISKLYWYQKPDGTVKLLIYHTRSLHSGVPS 82

Qy 80 RFGSGSGTDTFTTISLQPEDIAITYCOOYSLYRSFGGOTKVEIKRGGSGG--GSGS 136
Db 83 RFGSGSGTDTFTTISLQPEDIAITYCOOQNTLPTPTFGGCTKLEITGSTSGSKPGSGE 142

Qy 137 GGSSEVOLVESGGVQVQGRSLRSLSCSAGDFDTTYWMSWRQAPKGLWIGELHSDS 196
Db 143 GSTGKVLQSGFGLVAPQSLSVTCTVSGVSLFDYGVSWIROPPRKGLEWLVW--GS 201

Qy 197 STINVAFLKDRFTISRDNKNTLFLQMSLRPBDTGVYFCASLYF--GFPWFAYWGGT 254
Db 202 ETTYNGALKSRLLTIIDKNSKQVFLKNSLQTDITAIYCAKHYGGSVAMDIWGGT 261

Qy 255 PVTVSSAKPTTTPAPR--PPTPAPTASQ-----LSLRPEA----- 289
Db 262 SVTVSSVEPKSSDKTHCTPCPAPBELGGSPVFLPPKPKDTLMSRTPETVCVVVDVSH 321

Qy 290 -----AR 291
Db 322 EDPEVKFNWYVDGVEVHNNAKTPREOVNSTRYRVSVLTVLHQDWLNGKEYKCKVSNKAL 381

Qy 292 PA-----AGG-----AVHT-----RGL---DFALD-----PK 310
Db 382 PAPLEKTSKAGQPREPQVTVLPSPRDELTONQVSLTCLVKGFYPSDIAVEWESNGQPE 441

Qy 311 LCYLL-----DGILFIY----- 322
Db 442 NNYKTPVPLDSDGSFFLYSKLTVDKSRWQOQNVFSCSVMEALHNNHYTKSLSLSPGKM 501

Qy 323 -----GVIL---TALFLRVKFSRABPPAYOQGNOLYNELNLRREBEYDVKRR 370
Db 502 ALIVLGGVAGLLFIGLGIFPRVKFSRABPPAYOQGNOLYNELNLRREBEYDVKRR 561

Qy 371 GRDPEMGKPRKPNQEGLYNELQDKMAEYSEIGMKGERRRGKHGDLVYQGLSTATKD 430
Db 562 GRDPEMGKPRKPNQEGLYNELQDKMAEYSEIGMKGERRRGKHGDLVYQGLSTATKD 621

Qy 431 TYDALHMOALPPR 443

```

```

Db      622 TYDALHMOALPPR 634

RESULT 4
US-10-829-388-11
; Sequence 11, Application US/10829388
; Publication No. US20050003403A1
; GENERAL INFORMATION:
; APPLICANT: ROSSI, EDMUND A.
; APPLICANT: CHANG, CHIEN HSING
; APPLICANT: MCBRIDE, WILLIAM J.
; TITLE OF INVENTION: POLYVALENT PROTEIN COMPLEX
; FILE REFERENCE: 41133-0006US1
; CURRENT APPLICATION NUMBER: US/10/829,388
; PRIOR FILING DATE: 2004-04-22
; PRIOR APPLICATION NUMBER: 60/464,532
; PRIOR FILING DATE: 2003-04-22
; PRIOR APPLICATION NUMBER: 60/525,391
; PRIOR FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric sequence from multiple species
US-10-829-388-11

Query Match      51.3%; Score 1205.5; DB 16; Length 364;
Best Local Similarity 94.3%; Pred. No. 5.4e-73;
Matches 230; Conservative 0; Mismatches 5; Indels 9; Gaps 1;

Qy      16 GVHSDIQLTQSPSSLSASVGDRTVITCKASQDVGTSVAMYYQKPGKAPKLLIYWTSTRHT 75
Db      122 GGSDIQLTQSPSSLSASVGDRTVITCKASQDVGTSVAMYYQKPGKAPKLLIYWTSTRHT 181
Qy      76 GVPSRFGSGSGTDFTTISSLPEDIATYCCQYSLYRSFGQGTKEIKRGSGSGSG 135
Db      182 GVPSRFGSGSGTDFTTISSLPEDIATYCCQYSLYRSFGQGTKEIKR----- 232
Qy      136 SGGSGSEVOLVESGGVQGRSLRLSCSASGDFTTYMMSWVRQAPGKGLEWIGEIHDP 195
Db      233 LEGGSEVOLVESGGVQGRSLRLSCSASGDFTTYMMSWVRQAPGKGLEWIGEIHDP 292
Qy      196 SSTINYAPSLKDRFTTISRDNKNTLFLQMSLRPEDTGVYFCASLYFGFPWPAYWGQGP 255
Db      293 SSTINYAPSLKDRFTTISRDNKNTLFLQMSLRPEDTGVYFCASLYFGFPWPAYWGQGP 352
Qy      256 VTVS 259
Db      353 VTVS 356

RESULT 5
US-10-829-388-1
; Sequence 1, Application US/10829388
; Publication No. US20050003403A1
; GENERAL INFORMATION:
; APPLICANT: ROSSI, EDMUND A.
; APPLICANT: CHANG, CHIEN HSING
; APPLICANT: MCBRIDE, WILLIAM J.
; TITLE OF INVENTION: POLYVALENT PROTEIN COMPLEX
; FILE REFERENCE: 41133-0006US1
; CURRENT APPLICATION NUMBER: US/10/829,388
; PRIOR FILING DATE: 2004-04-22
; PRIOR APPLICATION NUMBER: 60/464,532
; PRIOR FILING DATE: 2003-04-22
; PRIOR APPLICATION NUMBER: 60/525,391
; PRIOR FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1

Query Match      51.3%; Score 1205.5; DB 16; Length 364;
Best Local Similarity 94.3%; Pred. No. 5.4e-73;
Matches 230; Conservative 0; Mismatches 5; Indels 9; Gaps 1;

Qy      16 GVHSDIQLTQSPSSLSASVGDRTVITCKASQDVGTSVAMYYQKPGKAPKLLIYWTSTRHT 75
Db      122 GGSDIQLTQSPSSLSASVGDRTVITCKASQDVGTSVAMYYQKPGKAPKLLIYWTSTRHT 181
Qy      76 GVPSRFGSGSGTDFTTISSLPEDIATYCCQYSLYRSFGQGTKEIKRGSGSGSG 135
Db      182 GVPSRFGSGSGTDFTTISSLPEDIATYCCQYSLYRSFGQGTKEIKR----- 232
Qy      136 SGGSGSEVOLVESGGVQGRSLRLSCSASGDFTTYMMSWVRQAPGKGLEWIGEIHDP 195
Db      233 LEGGSEVOLVESGGVQGRSLRLSCSASGDFTTYMMSWVRQAPGKGLEWIGEIHDP 292
Qy      196 SSTINYAPSLKDRFTTISRDNKNTLFLQMSLRPEDTGVYFCASLYFGFPWPAYWGQGP 255
Db      293 SSTINYAPSLKDRFTTISRDNKNTLFLQMSLRPEDTGVYFCASLYFGFPWPAYWGQGP 352
Qy      256 VTVS 259
Db      353 VTVS 356

RESULT 6
US-10-829-388-12
; Sequence 12, Application US/10829388
; Publication No. US20050003403A1
; GENERAL INFORMATION:
; APPLICANT: ROSSI, EDMUND A.
; APPLICANT: CHANG, CHIEN HSING
; APPLICANT: MCBRIDE, WILLIAM J.
; TITLE OF INVENTION: POLYVALENT PROTEIN COMPLEX
; FILE REFERENCE: 41133-0006US1
; CURRENT APPLICATION NUMBER: US/10/829,388
; PRIOR FILING DATE: 2004-04-22
; PRIOR APPLICATION NUMBER: 60/464,532
; PRIOR FILING DATE: 2003-04-22
; PRIOR APPLICATION NUMBER: 60/525,391
; PRIOR FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric sequence from multiple species
US-10-829-388-12

Query Match      51.2%; Score 1202.5; DB 16; Length 358;
Best Local Similarity 95.0%; Pred. No. 8.5e-73;
Matches 228; Conservative 0; Mismatches 3; Indels 9; Gaps 1;

Qy      20 DIQLTQSPSSLSASVGDRTVITCKASQDVGTSVAMYYQKPGKAPKLLIYWTSTRHTGVPS 79
Db      2 DIQLTQSPSSLSASVGDRTVITCKASQDVGTSVAMYYQKPGKAPKLLIYWTSTRHTGVPS 61
Qy      80 RFGSGSGTDFTTISSLPEDIATYCCQYSLYRSFGQGTKEIKRGSGSGSGSGS 139
Db      62 RFGSGSGTDFTTISSLPEDIATYCCQYSLYRSFGQGTKEIKRG-----GGQ 112
Qy      140 GSEVOLVESGGVQGRSLRLSCSASGDFTTYMMSWVRQAPGKGLEWIGEIHDPDSTI 199
Db      113 FMEVOLVESGGVQGRSLRLSCSASGDFTTYMMSWVRQAPGKGLEWIGEIHDPDSTI 172

```

OY 200 NYAPSLKDRFTISRDNAKNTLFLQMSLRPDTGVYFCASLYFGPPWYAYWGQGTPTVTS 259  
 DB 173 NYAPSLKDRFTISRDNAKNTLFLQMSLRPDTGVYFCASLYFGPPWYAYWGQGTPTVTS 232

## RESULT 7

US-10-829-388-2  
 ; Sequence 2, Application US/10829388  
 ; Publication No. US20050003403A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ROSSI, EDMUND A.  
 ; APPLICANT: CHANG, CHIEN HSING  
 ; APPLICANT: MCBRIDE, WILLIAM J.  
 ; TITLE OF INVENTION: POLYVALENT PROTEIN COMPLEX  
 ; FILE REFERENCE: 41133-0006051  
 ; CURRENT APPLICATION NUMBER: US/10/829,388  
 ; PRIORITY FILING DATE: 2004-04-22  
 ; PRIOR APPLICATION NUMBER: 60/464,532  
 ; PRIOR FILING DATE: 2003-04-22  
 ; PRIOR APPLICATION NUMBER: 60/525,391  
 ; PRIOR FILING DATE: 2003-11-24  
 ; NUMBER OF SEQ ID NOS: 20  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 2  
 ; LENGTH: 363  
 ; TYPE: PRP  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Chimeric sequence from multiple species  
 US-10-829-388-2

Query Match 51.2%; Score 1202.5; DB 16; Length 363;  
 Best Local Similarity 95.0%; Pred. No. 8.6e-73;  
 Matches 228; Conservative 0; Mismatches 3; Indels 9; Gaps 1;  
 OY 20 DIQLTQSPSSLSASVGDRTITCKASQDVGTISVAVYQKPGKAPKLLIYWTSTRHTGVPS 79  
 DB 7 DIQLTQSPSSLSASVGDRTITCKASQDVGTISVAVYQKPGKAPKLLIYWTSTRHTGVPS 66  
 OY 80 RFGSGSGTDTFTTSSLPEDATYTCQXSLYRSFGQTKVEIKRGSGSGSGSGS 139  
 DB 67 RFGSGSGTDTFTTSSLPEDATYTCQXSLYRSFGQTKVEIKRGSGSGSGSGS 117  
 OY 140 GSEVLVESGGVQVQGRSLRLSCASGDFFTYMSWVRQAPGKLEWIEIHPDSSTI 199  
 DB 118 FMEVLVESGGVQVQGRSLRLSCASGDFFTYMSWVRQAPGKLEWIEIHPDSSTI 177  
 OY 200 NYAPSLKDRFTISRDNAKNTLFLQMSLRPDTGVYFCASLYFGPPWYAYWGQGTPTVTS 259  
 DB 178 NYAPSLKDRFTISRDNAKNTLFLQMSLRPDTGVYFCASLYFGPPWYAYWGQGTPTVTS 237

## RESULT 8

US-08-812-393A-2  
 ; Sequence 2, Application US/08812393A  
 ; Publication No. US20010007152A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: SHERMAN, Linda A.  
 ; APPLICANT: LUSTGARTEN, Joseph  
 ; TITLE OF INVENTION: RECOMBINANT CONSTRUCTS ENCODING  
 ; TITLE OF INVENTION: T CELL RECEPTORS SPECIFIC FOR HUMAN HLA-RESTRICTED TUMOR  
 ; TITLE OF INVENTION: ANTIGENS  
 ; NUMBER OF SEQUENCES: 64  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: MORRISON & FOERSTER  
 ; STREET: 2000 Pennsylvania Avenue, NW, suite 5500  
 ; CITY: Washington  
 ; STATE: DC  
 ; COUNTRY: USA  
 ; ZIP: 20006-1888  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/812.393A  
 ; FILING DATE: 05-MAR-1997  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Murashige, Kate H  
 ; REGISTRATION NUMBER: 29,959  
 ; REFERENCE/DOCKET NUMBER: 31333-20001.00  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 202-887-1500  
 ; TELEFAX: 202-822-0168  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 444 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; FRAGMENT TYPE: internal  
 ; US-08-812-393A-2

Query Match 46.6%; Score 1094; DB 8; Length 444;  
 Best Local Similarity 54.7%; Pred. No. 2.1e-65;  
 Matches 248; Conservative 35; Mismatches 106; Indels 64; Gaps 12;  
 OY 22 QLTPSPSSLSASVGDRTITCKASQDVGTISVAVYQKPGKAPKLLIYWTSTRHTGVPSR 80  
 DB 25 QVQSPASLVLEGENAELQCSFS--IFTNQVQVFPQPGRLVSLLY-----NPSG 74  
 OY 81 FSGSGSGTDTFTTSSLPEDATYTCQOYS-----FTISSLPEDATYTCQOYS-----LYRSFGQTKVEIKRG 128  
 DB 75 TKQSGRLTSTTVIKERRSSLHSSQITDSTGYILCASNSGGSNAKLTFGKGTGLSVKSGG 134  
 OY 129 SGSGSGSGSGSEVQLVES--GGGVQVQGRSLRLSCASGDFFTYMSWVRQAPGKLE 187  
 DB 135 GSGSGSGSGSGSEAAVTQSPRNKAVTGGKVTLSCTQNNHNMV---WTRQDTGHLR 191  
 OY 188 WIGIHPDSSTINAPSLKDRFTISRDNAXN-TLFLQMSLRPDTGVYFCASLYFGPPW 246  
 DB 192 LIHYSYGAGST--EKGDIPDGYKASRPSQENFSLILELAT--PSQTSVYFCASGETGTNE 247  
 OY 247 PAYWQGTPTVTVSS-----AKPTTTPAPRPTPTAPTASQPLSLRPEAA 290  
 DB 248 RLFFGHGHTKLSVLTNSIMYFSHFVFPVFLPAKPTTTTAPRPTPTAPTASQPLSLRPS 307  
 OY 291 RPAAGGAVHTRGDLFDALDPKLCYLLDGLTIFYGVILTALFLRVKFSRAEPPAYQOQNG 350  
 DB 308 R-----DPKLCYLLDGLTIFYGVILTALFLRVKFSRAEPPAYQOQNG 351  
 OY 351 LYNELNLGRREYDVLDRGRDPEMGKPKRRKNPQEGLYNELQDKMAEAYSEIGMKE 410  
 DB 352 LYNELNLGRREYDVLDRGRDPEMGKPKRRKNPQEGLYNELQDKMAEAYSEIGMKE 411  
 OY 411 RRRKGHDGLYQGLSTATKQTYDALHMQALPPR 443  
 DB 412 RRRKGHDGLYQGLSTATKQTYDALHMQALPPR 444

## RESULT 9

US-09-774-681-2  
 ; Sequence 2, Application US/09774681  
 ; Publication No. US20030208780A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sunol Molecular Corporation  
 ; APPLICANT: Sherman, Linda  
 ; APPLICANT: Lustgarten, Joseph

```

; TITLE OF INVENTION: RECOMBINANT CONSTRUCTS ENCODING T CELL
; FILE REFERENCE: 31333-20001.01
; CURRENT APPLICATION NUMBER: US/09/774,681
; CURRENT FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: US 08/812,393
; PRIOR FILING DATE: 1997-03-05
; PRIOR APPLICATION NUMBER: US 60/012,845
; PRIOR FILING DATE: 1996-03-05
; SOFTWARE: FastSeq for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 65
; SEQ ID NO 2
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Deduced amino acid derivative of effective T cell
; OTHER INFORMATION: receptor
US-09-774-681-2

Query Match 46.5%; Score 1092; DB 10; Length 449;
Best Local Similarity 54.5%; Pred. No. 2.9e-65;
Matches 247; Conservative 36; Mismatches 106; Indels 64; Gaps 12;

Qy 22 QLTQSPSSLSASVGDRTVITCKASQDVGTGTS-VAVYQKPGKAPKLLIYWTSTRHTGVPGR 80
Db 25 QVQSPASLVQEGENAELOCSF--IFTNQVWFYQRPGRGLVSLLY-----NPSG 74
Qy 81 FSGSGSGTDTFT-----FTISLQPEDIATYCCQYS-----LYRSFGGQTKVEIKRG 128
Db 75 TKSGRLTSTTVIKERRSLHSSQITDSGTYLCASNGSGSNAKLTFGKTKLSVKGSG 134
Qy 129 SGSGSGSGSGSEVQLVES-GGVVQVQGRSLRLSCSASGDFDFTYMWVWVROAPGKGL 187
Db 135 GSGSGSGSGSGSEAAVTPQPRNKVAVTGKVLSCNQTNHNNY---WVRQDTGHGLR 191
Qy 188 WIGEIHDPDSSTINYAPSLKDRFTISRDNAXN-TLFLQMDSLRPEDTGVYFCASLYGFPW 246
Db 192 LIHYSYGAGST--EKGDIPDGKASRPSQENFSLIVELGT--PSQTSVYFCASGETGTNE 247
Qy 247 FAYWGQGTPTVTVSS-----AKPTTTPAPRPPTPAPTIASOPLSLRPEAA 290
Db 248 RLFPFGHGTKLSVLTNSIMYFSHFVPLPAKPTTTPAPRPPTPAPTIASOPLSLRPS 307
Qy 291 RPAAGGAVHTRGLDFALDPKLYLLDGLIFGYVILTALFLRVKFSRSAPPAQOQONQ 350
Db 308 R-----DPLKLYLLDGLIFGYVILTALFLRVKFSRSADAPAYQOQONQ 351
Qy 351 LYNELNLGRREYVDLDRGRDPEMGKPRRKNPQEGLYNELQKQKVAEYSEIGMKGE 410
Db 352 LYNELNLGRREYVDLDRGRDPEMGKPRRKNPQEGLYNELQKQKVAEYSEIGMKGE 411
Qy 411 RRRKGHDGLYQGLSTATKTYDALHMQALPPR 443
Db 412 RRRKGHDGLYQGLSTATKTYDALHMQALPPR 444

RESULT 10
US-09-203-958A-2
; Sequence 2, Application US/09203958A
; Publication No. US20030039641A1
; GENERAL INFORMATION:
; APPLICANT: KELLER, Tibor
; APPLICANT: GOLDSTEIN, Joel
; APPLICANT: GRAZIANO, Robert M.
; APPLICANT: DEO, Yashwant M.
; TITLE OF INVENTION: CELLS EXPRESSING ANTI-FC RECEPTOR
; TITLE OF INVENTION: BINDING COMPONENTS
; FILE REFERENCE: MXI-099CPA
; CURRENT APPLICATION NUMBER: US/09/203,958A
; CURRENT FILING DATE: 1998-12-02
; PRIOR APPLICATION NUMBER: 60/067232
; PRIOR FILING DATE: 1997-12-02

```

```

; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic construct
US-09-203-958A-2

Query Match 41.1%; Score 965; DB 10; Length 352;
Best Local Similarity 75.3%; Pred. No. 7.4e-57;
Matches 195; Conservative 13; Mismatches 43; Indels 8; Gaps 3;

Qy 12 ATATGVHSDIQLTQSPSSLSASVGDRTVITCKASQDVGTGTS-----VAVYQKPGKAPKL 65
Db 30 AGAQPARDIQLTQSPSSLSASVGDRTVITCKSSQSVLYSSNQKNVLAWYQKPGKAPKL 89
Qy 66 LIYWTSTRHTGVPGRFSGSGSGTDTFTTISLQPEDIATYCCQYSLYRSFGGQTKVEIK 125
Db 90 LIYWASTRESGVPSRFSFGSGSGTDTFTTISLQPEDIATYCHQYLSSWTFGGQTKVEIK 149
Qy 126 -RGSGSGSGSGSGSEVQLVESGGVQVQGRSLRLSCSASGDFDFTYMWVWVROAPGK 184
Db 150 SSGSGSGSGSGSGSEVQLVESGGVQVQGRSLRLSCSSSGFIISDNTYMWVWVROAPGK 209
Qy 185 GLEWIGEIHDPDSSTINYAPSLKDRFTISRDNAXNTLFLQMDSLRPEDTGVYFCASLYGFG 244
Db 210 GLEWVATISDGGSYTYTPDSVKGRFTISRDNAXNTLFLQMDSLRPEDTGVYFCARGYRY 269
Qy 245 PW-FAYWGQGTPTVTVSSAK 262
Db 270 EGAMDYWGQGTPTVTVSSPR 288

RESULT 11
US-10-120-198B-2
; Sequence 2, Application US/10120198B
; Publication No. US20030215427A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Michael
; TITLE OF INVENTION: CE7-SPECIFIC REDIRECTED IMMUNE CELLS
; FILE REFERENCE: 1954-337
; CURRENT APPLICATION NUMBER: US/10/120,198B
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/282,859
; PRIOR FILING DATE: 2001-04-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 631
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: mouse-human chimera
US-10-120-198B-2

Query Match 38.7%; Score 910; DB 15; Length 631;
Best Local Similarity 44.4%; Pred. No. 6.8e-53;
Matches 228; Conservative 47; Mismatches 118; Indels 120; Gaps 18;

Qy 16 GVHSDIQLTQSPSSLSASVGDRTVITCKASQDVGTGTSVAVYQKPGKAPKLLIYWTSTRHT 75
Db 154 GGSDIQLTQSPSSLSASVGDRTVITCKANEDINNLAWYQKPGKAPKLLIYWTSTRHT 213
Qy 76 GVPSRFSFGSGSGTDTFTTISLQPEDIATYCCQY-SLYRSFGGQTKVEIKRSGSGSGS 134
Db 214 GVPSRFSFGSGSGDKYTLTISLQAEAFATYCCQYQWSTPFTFGSGTELEIKVPEKSDKT 273
Qy 135 GSGSGSEVQLVESGGGV-----QGRSLRLSCSASGDFDFTYMWVWVROAPKGL 189
Db 274 HTCPPCPAPELL--GGPSVFLFPKPKDMLMISRTP---EVTGVVDVSHEDPEVKPNW- 327

```



```

RESULT 14
US-09-956-086-3
; Sequence 3, Application US/09956086
; Patent No. US20020155498A1
; GENERAL INFORMATION:
; APPLICANT: FILPULA, DAVID
; WANG, MAOLIANG
; SHORR, ROBERT
; WHITLOW, MARC
; LEE, LIHSYNG S.
; TITLE OF INVENTION: CAPABLE OF GLYCOSYLATION, PRODUCTION AND USES THEREOF
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNES, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/956,086
; FILING DATE: 20-Sep-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/069,821
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 60/063,074
; FILING DATE: 27-OCT-1997
; APPLICATION NUMBER: US 60/050,472
; FILING DATE: 23-JUN-1997
; APPLICATION NUMBER: US 60/044,449
; FILING DATE: 30-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: KIM, JUDITH U.
; REGISTRATION NUMBER: 40,679
; REFERENCE/DOCKET NUMBER: 0977.2280003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)371-2600
; TELEFAX: (202)371-2540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: No. US20020155498A1 Relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-956-086-3

Query Match 37.0%; Score 870; DB 9; Length 263;
Best Local Similarity 68.2%; Pred. No. 1.3e-50;
Matches 178; Conservative 23; Mismatches 42; Indels 18; Gaps 5;

Qy 20 DIQLTQSPSLASVGDVRVITTKASQ---DVGTSVANVYQKFGKAPKLLIYWTSTRTHTG 76
Db 1 DIQMTQSPSLASVGDVRVITTKASQSLVSIISNLAWSYQKFGKAPKLLIYAASLSLEG 60
Qy 77 VPSRFSGSGGTFTFTISSLPQEDATATYYCOOYSLYR--SFGQGTKEVTKRGSGSG-- 132
Db 61 VPSRFSGSGGTFTFTISSLPQEDFATYYCOQYNSLPFTFGQGTKEVTKGSTSGSKP 120
Qy 133 GSGSGSGSGSVQLVESGGGVVQPGRSRLRLSCSAGDFDTTYMNSWVRQAPGKGLEWIGE 192
Db 121 GSGEGSTKGEVQLVESGGGVVQPGGSLRLSCAASGFTTFSSYANSWVRQAPGKGLEWSVI 180
Qy 193 --HPDSTSTINYPSLKDRFTTISRDNKNTLFLQMDSLRPEDTGYFCA-----SLY 241

```

```

Db      181  SGTGDCGTYADSVKGRFTISRDNSKNTLYIQMNSLRADDTAVYYCARGKXGKSLSGXY 240
QY      242  FGPPWFAYWGQGTPTVTSSAK 262
       : : : |||||
Db      241  YYHYFDYWGQGLTIVTSSNK 261
       : : : |||||

RESULT 15
US-09-956-087-3
; Sequence 3, Application US/09956087
; Patent No. US20020161201A1
; GENERAL INFORMATION:
; APPLICANT: FILFULA, DAVID
;           WANG, MAOLIANG
;           SHORR, ROBERT
;           WHITLOW, MARC
;           LEE, LIHSYNG S.
; TITLE OF INVENTION: SINGLE-CHAIN ANTIGEN-BINDING PROTEINS
;                   CAPABLE OF GLYCOSYLATION, PRODUCTION AND USES THEREOF
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/956,087
; FILING DATE: 20-Sep-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/069,821
; FILING DATE: 1998-04-30
; APPLICATION NUMBER: US 60/063,074
; FILING DATE: 27-OCT-1997
; APPLICATION NUMBER: US 60/050,472
; FILING DATE: 23-JUN-1997
; APPLICATION NUMBER: US 60/044,449
; FILING DATE: 30-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: KIM, JUDITH U.
; REGISTRATION NUMBER: 40,679
; REFERENCE/DOCKET NUMBER: 0977.2280003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)371-2600
; TELEFAX: (202)371-2540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
;
US-09-956-087-3

Query Match      37.0%; Score 870; DB 9; Length 263;
Best Local Similarity 68.2%; Pred. No. 1.3e-50;
Matches 178; Conservative 23; Mismatches 42; Indels 18; Gaps 5;

QY      20  DIQLTQSPSSLSASVGDRTVITTKASQ---DVGTSVAVYQQKFGKAPKLLIYWTSTRHG 76
Db      1  DIQMTQSPSSLSASVGDRTVITTKASQSLVSNLYAWYQQKFGKAPKLLIYAASSLSLG 60
QY      77  VPSRFSGSGSGTDFTTITSLQPEDIATYYCQOYSLYR--SFQGTQKVEIKRGGSGSG-- 132

```

```

Db      181  SGTGCGSYTADSVKGRFTISRDNSKNTLYLQMNLSRAEDTAVYYCARGRXKXSLSGY 240
QY      242  FGFPWFAYWGQGTPTVTSSAK 262
Db      241  YYHYFDYWGQGLTIVTSSNK 261

RESULT 15
US-09-956-087-3
; Sequence 3, Application US/09956087
; Patent No. US20020161201A1
; GENERAL INFORMATION:
; APPLICANT: FILFULA, DAVID
; WANG, MAOLIANG
; SHORR, ROBERT
; WHITLOW, MARC
; LEE, LIHSYNG S.
; TITLE OF INVENTION: SINGLE-CHAIN ANTIGEN-BINDING PROTEINS
; CAPABLE OF GLYCOSYLATION, PRODUCTION AND USES THEREOF
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNER, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/956,087
; FILING DATE: 20-Sep-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/069,821
; FILING DATE: 1998-04-30
; APPLICATION NUMBER: US 60/063,074
; FILING DATE: 27-OCT-1997
; APPLICATION NUMBER: US 60/050,472
; FILING DATE: 23-JUN-1997
; APPLICATION NUMBER: US 60/044,449
; FILING DATE: 30-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: KIM, JUDITH U.
; REGISTRATION NUMBER: 40,679
; REFERENCE/DOCKET NUMBER: 0977.2280003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
;
US-09-956-087-3

Query Match      37.0%; Score 870; DB 9; Length 263;
Best Local Similarity 68.2%; Pred.No. 1.3e-50;
Matches 178; Conservative 23; Mismatches 42; Indels 18; Gaps 5;

QY      20  DIQLTSPSSLSASVGDVRTITTCASQ---DVGTSVAWYQQKFGKAPKLIYWTSTRHG 76
Db      1  DIQMTSPSSLSASVGDVRTITTCASQSLVSI SNLYAWYQQKFGKAPKLIYAASSLESG 60
QY      77  VPSRFSGSGSGTDFTTITSLQPEDTATYYCQSYLYR--SFGQGTKEVTKRGSGSG-- 132

```



